Practical Work in Geography

Part II

Textbook for Class XII

© NCERT not to be republished

12101

राष्ट्रीय शैक्षिक अनुसंधान और प्रशिक्षण परिषद्
NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING

Reprint 2026-27

12101 – Practical Work in Geography

Textbook for Class XII

ISBN 81-7450-702-7

First Edition
February 2007 Magha 1928
Reprinted
February 2008 Magha 1929
January 2009 Magha 1930
January 2010 Magha 1931
December 2010 Pausa 1932
January 2014 Magha 1935
February 2015 Magha 1936
December 2015 Agrahayana 1937
February 2017 Magha 1938
December 2017 Pausa 1939
December 2018 Agrahayana 1940
August 2019 Shravana 1941
January 2021 Pausa 1942
December 2021 Agrahayana 1943

Revised Edition
August 2022 Sharvana 1944

Reprinted
March 2024 Chaitra 1946
December 2024 Agrahayana 1946
February 2026 Phalguna 1947

PD 20T HK

© National Council of Educational Research and Training, 2007, 2022

ALL RIGHTS RESERVED

- No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of the publisher.
- This book is sold subject to the condition that it shall not, by way of trade, be lent, re-sold, hired out or otherwise disposed of without the publisher's consent, in any form of binding or cover other than that in which it is published.
- The correct price of this publication is the price printed on this page, Any revised price indicated by a rubber stamp or by a sticker or by any other means is incorrect and should be unacceptable.

OFFICES OF THE PUBLICATION DIVISION, NCERT

NCERT Campus
Sri Aurobindo Marg
New Delhi 110 016 **Phone : 011-26562708**

108, 100 Feet Road
Hosdakere Halli Extension
Banashankari III Stage
Bengaluru 560 085 **Phone : 080-26725740**

Navjivan Trust Building
P.O.Navjivan
Ahmedabad 380 014 **Phone : 079-27541446**

CWC Campus
Opp. Dhankal Bus Stop
Panihati
Kolkata 700 114 **Phone : 033-25530454**

CWC Complex
Maligaon
Guwahati 781 021 **Phone : 0361-2674869**

Publication Team

Head, Publication Division	:	*M.V. Srinivasan*
Chief Editor	:	*Bijnan Sutar*
Chief Business Manager	:	*Amitabh Kumar*
Chief Production Officer (In charge)	:	*Deepak Jaiswal*
Editor	:	*M.G. Bhagat*
Production Assistant	:	*Sahitya Sharma*

Cover and Layout
Blue Fish

Cartography
Cartographic Design Agency

₹ 55.00

Printed on 80 GSM paper with NCERT watermark

Published at the Publication Division by the Secretary, National Council of Educational Research and Training, Sri Aurobindo Marg, New Delhi 110 016 and printed at SRG Traders Pvt. Ltd., B-41, Sector-67, Noida - 201 301 (U.P.)

© NCERT not to be republished

Foreword

The National Curriculum Framework (NCF), 2005, recommends that children's life at school must be linked to their life outside the school. This principle marks a departure from the legacy of bookish learning which continues to shape our system and causes a gap between the school, home and community. The syllabi and textbooks developed on the basis of NCF signify an attempt to implement this basic idea. They also attempt to discourage rote learning and the maintenance of sharp boundaries between different subject areas. We hope these measures will take us significantly further in the direction of a child-centred system of education outlined in the National Policy on Education (1986).

The success of this effort depends on the steps that school principals and teachers will take to encourage children to reflect on their own learning and to pursue imaginative activities and questions. We must recognise that, given space, time and freedom, children generate new knowledge by engaging with the information passed on to them by adults. Treating the prescribed textbook as the sole basis of examination is one of the key reasons why other resources and sites of learning are ignored. Inculcating creativity and initiative is possible if we perceive and treat children as participants in learning, not as receivers of a fixed body of knowledge.

These aims imply considerable change in school routines and mode of functioning. Flexibility in the daily time-table is as necessary as rigour in implementing the annual calendar so that the required number of teaching days are actually devoted to teaching. The methods used for teaching and evaluation will also determine how effective this textbook proves for making children's life at school a happy experience, rather than a source of stress or boredom. Syllabus designers have tried to address the problem of curricular burden by restructuring and reorienting knowledge at different stages with greater consideration for child psychology and the time available for teaching. The textbook attempts to enhance this endeavour by giving higher priority and space to opportunities for contemplation and wondering, discussion in small groups, and activities requiring hands-on experience.

The National Council of Educational Research and Training (NCERT) appreciates the hard work done by the textbook development committee responsible for this book. We wish to thank the Chairperson of the advisory committee for textbooks in Social Sciences, at the higher secondary level, Professor Hari Vasudevan and the Chief Advisor for this book, Professor M.H. Qureshi for guiding the work of this committee. Several teachers contributed to the development of this textbook; we are grateful to their principals for making this possible. We are indebted to the institutions and organisations which have

© NCERT not to be republished

generously permitted us to draw upon their resources, material and personnel. We are especially grateful to the members of the National Monitoring Committee, appointed by the Department of Secondary and Higher Education, Ministry of Human Resource Development under the Chairpersonship of Professor Mrinal Miri and Professor G.P. Deshpande, for their valuable time and contribution. As an organisation committed to systemic reform and continuous improvement in the quality of its products, NCERT welcomes comments and suggestions which will enable us to undertake further revision and refinement.

Director
National Council of Educational
Research and Training

New Delhi
20 November 2006

© NCERT not to be republished

Rationalisation of Content in the Textbooks

In view of the COVID-19 pandemic, it is imperative to reduce content load on students. The National Education Policy 2020, also emphasises reducing the content load and providing opportunities for experiential learning with creative mindset. In this background, the NCERT has undertaken the exercise to rationalise the textbooks across all classes. Learning Outcomes already developed by the NCERT across classes have been taken into consideration in this exercise.

Contents of the textbooks have been rationalised in view of the following:

- Overlapping with similar content included in other subject areas in the same class
- Similar content included in the lower or higher class in the same subject
- Difficulty level
- Content, which is easily accessible to students without much interventions from teachers and can be learned by children through self-learning or peer-learning
- Content, which is irrelevant in the present context

This present edition, is a reformatted version after carrying out the changes given above.

© NCERT not to be republished

© NCERT
not to be republished

Textbook Development Committee

Chairperson, Advisory Committee for Textbooks in Social Sciences at the Higher Secondary Level

Hari Vasudevan, *Professor,* Department of History, University of Calcutta, Kolkata

Chief Advisor

M. H. Qureshi, *Professor,* Centre for the Study of Regional Development, Jawaharlal Nehru University, New Delhi

Advisor

S. M. Rashid, *Professor*, Jamia Millia Islamia, New Delhi

Members

K. K. Sharma, *Principal (Retd.),* Lohia College, Churu

M. H. Quasmi, *Lecturer,* IASE, Jamia Millia Islamia, New Delhi

R. N. Vyas, *Professor,* CSSH, Mohanlal Sukhadia University, Udaipur

Shahab Fazal, *Reader,* Aligarh Muslim University, Aligarh

Sucharita Sen, *Associate Professor*, CSRD, Jawaharlal Nehru University, New Delhi

Member-Coordinator

Tannu Malik, *Lecturer*, Department of Education in Social Sciences and Humanities, NCERT, New Delhi

© NCERT not to be republished

Reprint 2026-27

Nirmalya Chakraborty, College of Art, New Delhi

© NCERT not to be republished

Acknowledgements

The National Council of Educational Research and Training acknowledges the contributions of H. Ramachandran, *Professor and Head,* Delhi School of Economics, Delhi University; B. S. Butola, *Professor,* CSRD, JNU; Odilia Coutinho, *Reader,* R.P.D. College, Belgaum; Anup Saikia, *Reader,* Gauhati University, Guwahati; Abdul Shaban, *Asstt. Professor,* Tata Institute of Social Sciences, Mumbai and Rupa Das, *PGT,* DPS, R.K. Puram, New Delhi in the development of this textbook.

Special thanks are due to Savita Sinha, *Professor and Head,* Department of Education in Social Sciences and Humanities for her valuable support at every stage of preparation of this textbook.

The Council is thankful to the Survey of India for certification of maps given in the textbook. It also gratefully acknowledges the support of individuals and organisations as listed below for providing various photographs and illustrations used in this textbook:

S.M. Rashid, *Professor,* Jamia Millia Islamia, New Delhi for fig. 1.2, 1.3 and 1.4; M.H. Quasmi, *Lecturer,* IASE, Jamia Millia Islamia, New Delhi for fig. 3.9, 3.10, 3.11 and 3.12; and Shahab Fazal, *Reader,* Aligarh Muslim University, Aligarh for fig. 4.8, 4.9, 4.10, 4.12 and 4.13.

The Council also gratefully acknowledges the contribution of Anil Sharma and Ishwar Singh *DTP Operators*; Ajay Singh, *Copy Editor,* Aarati Baloni, *Proof Reader* and Dinesh Kumar, *Computer Incharge* who have helped in giving a final shape to this book. The contribution of the Publication Department, NCERT in bringing out this textbook is also duly acknowledged.

© NCERT not to be republished

The following are applicable to all the maps of India used in this textbook

1. © Government of India, Copyright 2006
2. The responsibility for the correctness of internal details rests with the publisher.
3. The territorial waters of India extend into the sea to a distance of twelve nautical miles measured from the appropriate base line.
4. The administrative headquarters of Chandigarh, Haryana and Punjab are at Chandigarh.
5. The interstate boundaries amongst Arunachal Pradesh, Assam and Meghalaya shown on this map are as interpreted from the "North-Eastern Areas (Reorganisation) Act.1971," but have yet to be verified.
6. The external boundaries and coastlines of India agree with the Record/Master Copy certified by Survey of India.
7. The state boundaries between Uttaranchal and Uttar Pradesh, Bihar and Jharkhand, Chhattisgarh and Madhya Pradesh have not been verified by the Governments concerned.
8. The spellings of names in this map, have been taken from various sources.

© NCERT not to be republished

Reprint 2026-27

Contents

FOREWORD *iii*

CHAPTER 1

Data – Its Source and Compilation 1 – 12

CHAPTER 2

Data Processing 13 – 22

CHAPTER 3

Graphical Representation of Data 23– 45

CHAPTER 4

Spatial Information Technology 46 – 61

GLOSSARY 62

© NCERT not to be republished

THE CONSTITUTION OF INDIA

PREAMBLE

WE, THE PEOPLE OF INDIA, having solemnly resolved to constitute India into a [1]**[SOVEREIGN SOCIALIST SECULAR DEMOCRATIC REPUBLIC]** and to secure to all its citizens :

JUSTICE, social, economic and political;

LIBERTY of thought, expression, belief, faith and worship;

EQUALITY of status and of opportunity; and to promote among them all

FRATERNITY assuring the dignity of the individual and the [2][unity and integrity of the Nation];

IN OUR CONSTITUENT ASSEMBLY this twenty-sixth day of November, 1949 do **HEREBY ADOPT, ENACT AND GIVE TO OURSELVES THIS CONSTITUTION.**

---

1. Subs. by the Constitution (Forty-second Amendment) Act, 1976, Sec.2, for "Sovereign Democratic Republic" (w.e.f. 3.1.1977)
2. Subs. by the Constitution (Forty-second Amendment) Act, 1976, Sec.2, for "Unity of the Nation" (w.e.f. 3.1.1977)

12101CH01

1 Data – Its Source and Compilation

You must have seen and used various forms of data. For example, at the end of almost every news bulletin on Television, the temperatures recorded on that day in major cities are displayed. Similarly, the books on the Geography of India show data relating to the growth and distribution of population, and the production, distribution and trade of various crops, minerals and industrial products in tabular form. Have you ever thought what they mean? From where these data are obtained? How are they tabulated and processed to extract meaningful information from them ? In this chapter, we will deliberate on these aspects of the data and try to answer these many questions.

What is Data?

The data are defined as numbers that represent measurements from the real world. **Datum** is a single measurement. We often read the news like 20 centimetres of continuous rain in Barmer or 35 centimetres of rain at a stretch in Banswara in 24 hours or information such as New Delhi – Mumbai distance via Kota – Vadodara is 1385 kilometres and via Itarsi - Manmad is 1542 kilometres by train. This numerical information is called data. It may be easily realised that there are large volume of data available around the world today. However, at times, it becomes difficult to derive logical conclusions from these data if they are in raw form. Hence, it is important to ensure that the measured information is algorithmically derived and/or logically deduced and/or statistically calculated from multiple data. **Information** is defined as either a meaningful answer to a query or a meaningful stimulus that can cascade into further queries.

Need of Data

Maps are important tools in studying geography. Besides, the distribution and growth of phenomena are also explained through the data in tabular form. We know that an interelationship exists between many phenomena over the surface of the earth. These interactions are influenced by many variables which can be

© NCERT not to be republished

explained best in quantitative terms. Statistical analysis of those variables has become a necessity today. For example, to study cropping pattern of an area, it is necessary to have statistical information about the cropped area, crop yield and production, irrigated area, amount of rainfall and inputs like use of fertiliser, insecticides, pesticides, etc. Similarly, data related to the total population, density, number of migrants, occupation of people, their salaries, industries, means of transportation and communication is needed to study the growth of a city. Thus, data plays an important role in geographical analysis.

Presentation of the Data

You might have heard the story of a person who was travelling with his wife and a five-year old child. On his way, he had to cross a river. Firstly, he fathomed the depth of the river at four points as 0.6, 0.8, 0.9 and 1.5 metres. He calculated the average depth as 0.95 metres. His child's height was 1 metre. So, he led them to cross the river and his child drowned in the river. On the other bank, he sat pondering: "*Lekha Jokha Thahe, to Bachha Dooba Kahe*?" (Why did the child drown when average depth was within the reach of each one?). This is called statistical fallacy, which may deviate you from the real situation. So, it is important to collect the data to know the facts and figures, but equally important is the presentation of data. Today, the use of statistical methods in the analysis, presentation and in drawing conclusions plays a significant role in almost all disciplines, including geography, which use the data. It may, therefore, be inferred that the concentration of a phenomenon, e.g., population, forest or network of transportation or communication not only vary over space and time but may also be conveniently explained using the data. In other words, you may say that there is a shift from qualitative description to quantitative analysis in explaining the relationship among variables. Hence, analytical tools and techniques have become more important these days to make the study more logical and derive precise conclusion. Precise quantitative techniques are used right from the beginning of collecting and compiling data to its tabulation, organisation, ordering and analysis till the derivation of conclusions.

Sources of Data

The data are collected through the following ways. These are : 1. Primary Sources, and 2. Secondary Sources.

The data which are collected for the first time by an individual or the group of individuals, institution/organisations are called **Primary sources of the data**. On the other hand, data collected from any published or unpublished sources are called **Secondary sources**. *Fig. 1.1* shows the different methods of data collection.

Sources of Primary Data

1. Personal Observations

It refers to the collection of information by an individual or group of individuals through direct observations in the field. Through a field survey, information about the relief features, drainage patterns, types of soil and natural vegetation, as well as, population structure, sex ratio, literacy, means of transport and communication, urban and rural settlements, etc., is collected. However, in

© NCERT not to be republished

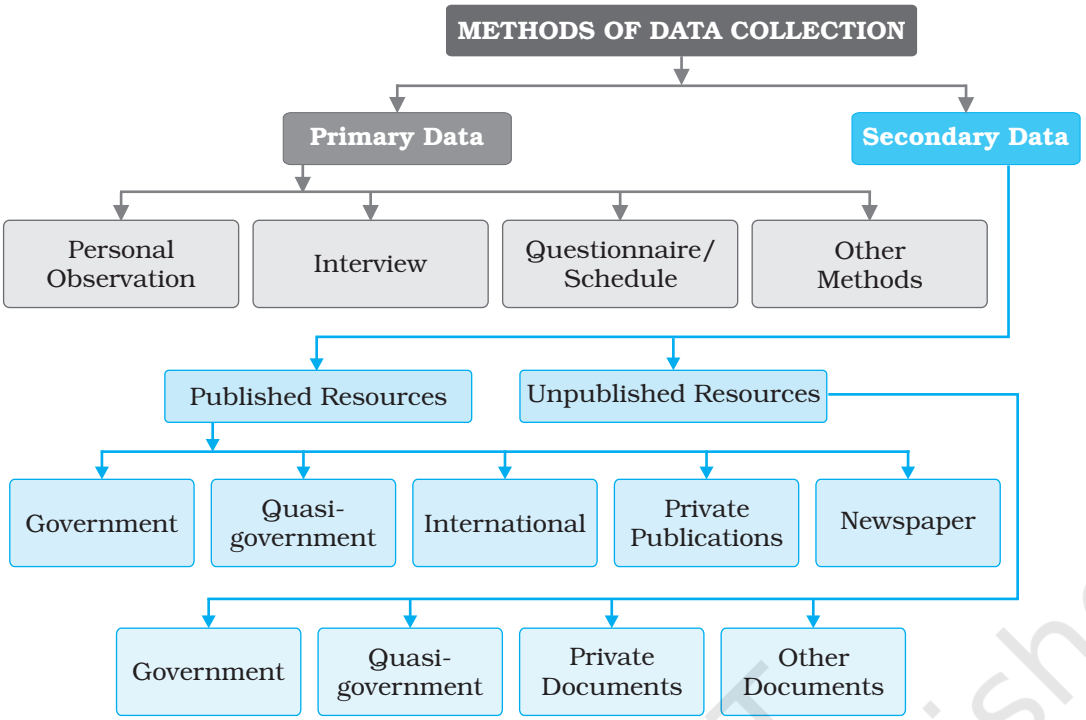


Fig. 1.1 : *Methods of Data Collection*

carrying out personal observations, the person(s) involved must have theoretical knowledge of the subject and scientific attitude for unbiased evaluation.

2. Interview

In this method, the researcher gets direct information from the respondent through dialogues and conversations. However, the interviewer must take the following precautions while conducting an interview with people of the area:

(i) A precise list of items about which information is to be gathered from the persons interviewed be prepared.

(ii) The person(s) involved in conducting the interview should be clear about the objective of the survey.

(iii) The respondents should be taken into confidence before asking any sensitive question and he/she be assured that the secrecy will be maintained.

(iv) A congenial atmosphere should be created so that the respondent may explain the facts without any hesitation.

(v) The language of the questions should be simple and polite so that the respondents feel motivated and readily agree to give the information asked for.

(vi) Avoid asking any such question that may hurt the self-respect or the religious feelings of the respondent.

(vii) At the end of the interview, ask the respondent what additional information he/she may provide, other than what has already been provided by him/her.

(viii) Pay your thanks and gratefulness for sparing his/her valuable time for you.

© NCERT not to be republished

3. Questionnaire/Schedule

In this method, simple questions and their possible answers are written on a plain paper and the respondents have to tick-mark the possible answers from the given choices. At times, a set of structured questions are written and sufficient space is provided in the questionnaire where the respondent write their opinion. The objectives of the survey should be clearly mentioned in the questionnaire. This method is useful in carrying out the survey of a larger area. Even questionnaire can be mailed to far-flung places. The limitation of the method is that only the literate and educated people can be approached to provide the required information. Similar to the questionnaire that contains the questions pertaining to the matter of investigation is the **schedule**. The only difference between the **questionnaire** and the **schedule** is that the respondent himself/herself fills up the questionnaires, whereas, a properly trained enumerator himself fills up schedules by asking question addressed to the respondents. The main advantage of schedule over the questionnaire is that the information from both literate and illiterate respondents can be collected.

4. Other Methods

The data about the properties of soil and water are collected directly in the field by measuring their characteristics using soil kit and water quality kit. Similarly, field scientists collect data about the health of the crops and vegetation using transducers (*Fig. 1.2*).

Fig. 1.2 : *Field Scientist taking Measures of Crop Health*

Secondary Source of Data

Secondary sources of data consist of published and unpublished records which include government publications, documents and reports.

Published Sources

1. Government Publications

The publications of the various ministries and the departments of the Government of India, state governments and the District Bulletins are one of the most important sources of secondary information. These include the Census of India published by the Office of the Registrar General of India, reports of the National Sample Survey, Weather Reports of Indian Meteorological Department and Statistical Abstracts published by state governments, and the periodical reports published by different Commissions. Some of the government publications are shown in *Fig. 1.3*.

Fig. 1.3 : *Some of the Government Publications*

© NCERT not to be republished

2. Semi/Quasi-government Publications

The publications and reports of Urban Development Authorities and Municipal Corporations of various cities and towns, Zila Parishads (District Councils), etc. fall under this category.

3. International Publications

The international publications comprise yearbooks, reports and monographs published by different agencies of the United Nations, such as United Nations Educational, Scientific and Cultural Organisation (UNESCO), United Nations Development Programme (UNDP), World Health Organisation (WHO), Food and Agricultural Organisation (FAO), etc. Some of the important publications of the United Nations that are periodically published are Demographic Year Book, Statistical Year Book and the Human Development Report (*Fig. 1.4*).

Fig. 1.4 : *Some of the United Nations Publications*

4. Private Publications

The yearbooks, surveys, research reports and monographs published by newspapers and private organisations fall under this category.

5. Newspapers and Magazines

The daily newspapers and the weekly, fortnightly and monthly magazines serve as easily accessible sources of secondary data.

6. Electronic Media

The electronic media, specially the internet, has emerged as a major source of secondary data in recent times.

Unpublished Sources

1. Government Documents

The unpublished reports, monographs and documents are yet another source of secondary data. These documents are prepared and maintained as unpublished record at different levels of governance. For example, the village level revenue records maintained by the *patwari* of respective villages serve as an important source of village-level information.

2. Quasi-government Records

The periodical reports and the development plans prepared and maintained by different Municipal Corporations, District Councils and Civil Services departments are included in Quasi-government records.

© NCERT not to be republished

3. Private Documents

These include unpublished reports and records of companies, trade unions, different political and apolitical organisations and residents' welfare associations.

Tabulation and Classification of Data

The data collected from primary or secondary sources initially appear as a big jumble of information with the least of comprehension. This is known as raw data. To draw meaningful inferences and to make them usable the raw data requires tabulation and classification.

One of the simplest devices to summarise and present the data is the **Statistical Table**. It is a systematic arrangement of data in columns and rows. The purpose of table is to simplify the presentation and to facilitate comparisons. This table enables the reader to locate the desired information quickly. Thus, the tables make it possible for the analyst to present a huge mass of data in an orderly manner within a minimum of space.

Data Compilation and Presentation

Data are collected, tabulated and presented in a tabular form either in absolute terms, percentages or indices.

Absolute Data

When data are presented in their original form as integers, they are called absolute data or **raw data**. For example, the total population of a country or a state, the total production of a crop or a manufacturing industry, etc. *Table 1.1* shows the absolute data of population of India and some of the selected states.

Table 1.1 : Population of India and Selected States/Union Territories, 2011

State/ UT Code	India/State/ Union Territory	Total Population		
		Persons	Males	Females
1	2	3	4	5
	INDIA[1]	**1,21,05,69,573**	**62,31,21,843**	**58,74,47,730**
1.	Jammu and Kashmir[2]	1,25,41,302	66,40,662	59,00,640
2.	Himachal Pradesh	68,64,602	34,81,873	33,82,729
3.	Punjab	2,77,43,338	1,46,39,465	1,31,03,873
4.	Chandigarh[3]	10,55,450	5,80,663	4,74,787
5.	Uttarakhand	1,00,86,292	51,37,773	49,48,519
6.	Haryana	2,53,51,462	1,34,94,734	1,18,56,728
7.	National Capital Territory of Delhi	1,67,87,941	89,87,326	78,00,615
8.	Rajasthan	6,85,48,437	3,55,50,997	3,29,97,440
9.	Uttar Pradesh	19,98,12,341	10,44,80,510	9,53,31,831
10.	Bihar	10,40,99,452	5,42,78,157	4,98,21,295

[1] inclusive of all territorial boundary of India
[2] excluding PoK
[3] Union Territory

Source : Census, 2011

Percentage/Ratio

Some time data are tabulated in a ratio or percentage form that are computed from a common parameter, such as literacy rate or growth rate of population, percentage of agricultural products or industrial products, etc. *Table 1.2* presents

© NCERT not to be republished

literacy rates of India over the decades in a percentage form. Literacy rate is calculated as :

$$\frac{\text{Total Literates}}{\text{Total Population}} \times 100$$

Table 1.2 : Literacy Rate : 1951 – 2011

Year	Person	Male	Female
1951	18.33	27.16	8.86
1961	28.3	40.4	15.35
1971	34.45	45.96	21.97
1981	43.57	56.38	29.76
1991	52.21	64.13	39.29
2001	64.84	75.85	54.16
2011	73.0	80.9	64.6

Source: Census, 2011

Index Number

An index number is a statistical measure designed to show changes in variable or a group of related variables with respect to time, geographic location or other characteristics. It is to be noted that index numbers not only measure changes over a period of time but also compare economic conditions of different locations, industries, cities or countries. Index number is widely used in economics and business to see changes in price and quantity. There are various methods for the calculation of index number. However, the simple aggregate method is most commonly used. It is obtained using the following formula:

$$\frac{\sum q_1}{\sum q_0} \times 100$$

$\sum q_1$ = Total of the current year production

$\sum q_0$ = Total of the base year production

Generally, base year values are taken as 100 and index number is calculated thereupon. For example, *Table 1.3* shows the production of iron ore in India and the changes in index number from 1970–71 to 2000–01 taking 1970–71 as the base year.

Table 1.3 : Production of Iron Ore in India

	Production (in million tonnes)	Calculation	Index Number
1970-71	32.5	$\frac{32.5}{32.5} \times 100$	100
1980-81	42.2	$\frac{42.2}{32.5} \times 100$	130
1990-91	53.7	$\frac{53.7}{32.5} \times 100$	165
2000-01	67.4	$\frac{67.4}{32.5} \times 100$	207

Source – India: Economic Year Book, 2005

Processing of Data

The processing of raw data requires their tabulation and classification in selected classes. For example, the data given in *Table 1.4* can be used to understand how they are processed.

We can see that the given data are ungrouped. Hence, the first step is to group data in order to reduce its volume and make it easy to understand.

© NCERT not to be republished

Table 1.4 : Score of 60 Students in Geography Paper

47	02	39	64	22	46	28	02	09	10
89	96	74	06	26	15	92	84	84	90
32	22	53	62	73	57	37	44	67	50
18	51	36	58	28	65	63	59	75	70
56	58	43	74	64	12	35	42	68	80
64	37	17	31	41	71	56	83	59	90

Grouping of Data

The grouping of the raw data requires determining of the number of classes in which the raw data are to be grouped and what will be the class intervals. The selection of the class interval and the number of classes, however, depends upon the range of raw data. The raw data given in *Table 1.4* ranges from 02 to 96. We can, therefore, conveniently choose to group the data into ten classes with an interval of ten units in each group, e.g. 0–10, 10–20, 20–30, etc. (*Table 1.5*).

Table 1.5 : Making Tally Marks to Obtain Frequency

Group	*Numerical of Raw Data*	*Tally Marks*	*Number of Individual*
0-10	02,02,09,06	////	4
10-20	10,15,18,12,17	𝍸	5
20-30	22,28,26,22,28	𝍸	5
30-40	39,32,37,36,35,37,31	𝍸//	7
40-50	47,46,44,43,42,41	𝍸/	6
50-60	53,57,50,51,58, 59,56,58,56,59	𝍸 𝍸	10
60-70	64,62,67,65, 63,64,68,64	𝍸 ///	8
70-80	74,73,75,70,74,71	𝍸 /	6
80-90	89,84,84,80,83	𝍸	5
90-100	96,92,90,90	////	4
			$\sum f = N = 60$

Process of Classification

Once the number of groups and the class interval of each group are determined, the raw data are classified as shown in *Table 1.5*. It is done by a method popularly known as **Four and Cross Method** or tally marks.

First of all, one tally mark is assigned to each individual in the group in which it is falling. For example, the first numerical in the raw data is 47. Since, it falls in the group of 40–50, one tally mark is recorded in the column 3 of *Table 1.5*.

Frequency Distribution

In *Table 1.5* we have classified the raw data of a quantitative variable and have grouped them class-wise. The number of individuals (places in the fourth column of *Table 1.5*) is known as frequency and the column represents the frequency

© NCERT not to be republished

distribution. It illustrates how the different values of a variable are distributed in different classes. Frequencies are classified as **Simple** and **Cumulative frequencies**.

Simple Frequencies

It is expressed by **'*f*'** and represent the number of individuals falling in each group *(Table 1.6)*. The sum of all the frequencies, assigned to all classes, represents the total number of individual observations in the given series. In statistics, it is expressed by the symbol N that is equal to $\sum f$. It is expressed as $\sum f = N = 60$ (*Table 1.5* and *1.6*).

Table 1.6 : Frequency Distribution

Group	*f*	*Cf*
00-10	4	4
10-20	5	9
20-30	5	14
30-40	7	21
40-50	6	27
50-60	10	37
60-70	8	45
70-80	6	51
80-90	5	56
90-100	4	60
	$\sum f = N = 60$	

Cumulative Frequencies

It is expressed by **'*Cf*'** and can be obtained by adding successive simple frequencies in each group with the previous sum, as shown in the column 3 of *Table 1.6*. For example, the first simple frequency in *Table 1.6* is 4. Next frequency of 5 is added to 4 which gives a total of 9 as the next cumulative frequency. Likewise, add every next number until the last cumulative frequency of 60 is obtained. Note that it is equal to N or $\sum f$.

Advantage of cumulative frequency is that one can easily make out that there are 27 individuals scoring less than 50 or that 45 out of 60 individuals lie below the score of 70.

Each simple frequency is associated with its group or class. The **exclusive** or **inclusive** methods are used for forming the groups or classes.

Exclusive Method

As shown in *Table 1.6*, two numbers are shown in its first column . Notice that the upper limit of one group is the same as the lower limit of the next group. For example, the upper limit of the one group (20 – 30) is 30, which is the lower limit of the next group (30 – 40), making 30 to appear in both groups. But any observation having the value of 30 is included in the group where it is at its lower limit and it is excluded from the group where it is the upper limit as (in 20-30 groups). That is why the method is known as exclusive method, i.e. a group is excluded of its upper limits. You may now make out where all the marginal values of *Table 1.4* will go.

The groups in *Table 1.6*, are interpreted in the following manner –

0 and under 10	10 and under 20
20 and under 30	30 and under 40
40 and under 50	50 and under 60
60 and under 70	70 and under 80
80 and under 90	90 and under100

Hence, in this type of grouping the class extends over ten units. For example, 20, 21, 22, 23, 24, 25, 26, 27, 28 and 29 are included in the third group.

© NCERT not to be republished

Inclusive Method

In this method, a value equal to the upper limit of a group is included in the same group. Therefore, it is known as inclusive method. Classes are mentioned in a different form in this method, as shown in the first column of *Table 1.7*. Normally, the upper limit of a group differs by 1 with the lower limits of the next group. It is important to note that each group spreads over ten units in this method also. For example, the group of 50–59 includes the ten values i.e. 50, 51, 52, 53, 54, 55, 56, 57, 58 and 59 (*Table 1.7*). In this method, both the upper and lower limit are included to find the frequency distribution.

Table 1.7 : Frequency Distribution

Group	f	Cf
0 – 9	4	4
10 – 19	5	9
20 – 29	5	14
30 – 39	7	21
40 – 49	6	27
50 – 59	10	37
60 – 69	8	45
70 – 79	6	51
80 – 89	5	56
90 – 99	4	60
	$\sum f = N = 60$	

Frequency Polygon

A graph of frequency distribution is known as the frequency polygon. It helps in comparing two or more than two frequency distributions (*Fig.1.5*). The two frequencies are shown using a bar diagram and a line graph respectively.

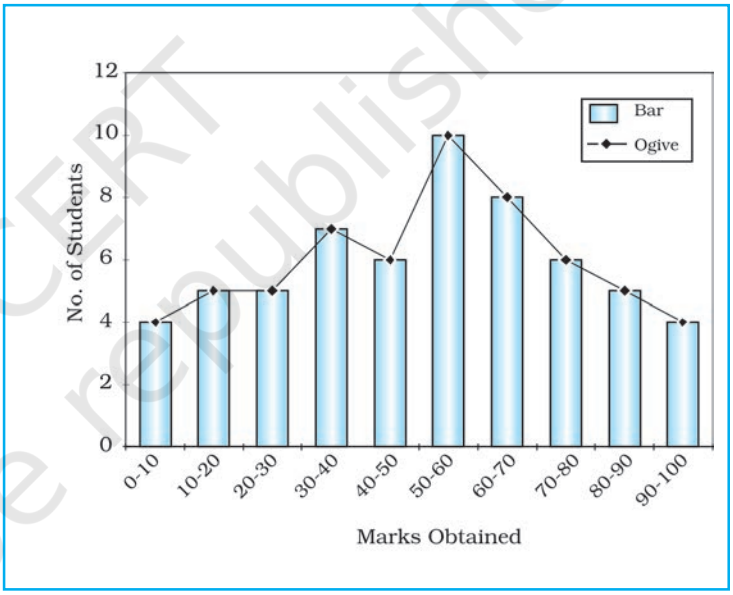


Fig. 1.5 : *Frequency Distribution Polygon*

Ogive

When the frequencies are added they are called cumulative frequencies and are listed in a table called cumulative frequency table. The curve obtained by plotting cumulative frequencies is called an **Ogive** (pronounced as ojive). It is constructed either by the **less than method** or the **more than method**.

In the **less than method**, we start with the upper limit of the classes and go on adding the frequencies. When these frequencies are plotted, we get a rising curve as shown in *Table 1.8* and *Fig. 1.6*.

In the **more than method,** we start with the lower limits of the classes and from the cumulative frequency, we substract frequency of each class. When these frequencies are plotted, we get a declining curve as shown in *Table 1.9* and *Fig 1.7*.

Both the *Figs. 1.5* and *1.6* may be combined to get a comparative picture of less than and more than Ogive as shown in *Table 1.10* and *Fig. 1.7*.

© NCERT not to be republished

Table 1.8 : Frequency Distribution less than Method

Less than Method	Cf
Less than 10	4
Less than 20	9
Less than 30	14
Less than 40	21
Less than 50	27
Less than 60	37
Less than 70	45
Less than 80	51
Less than 90	56
Less than 100	60

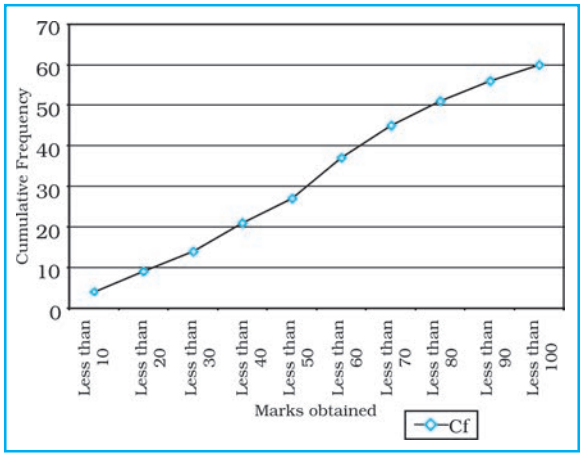


Fig. 1.6 : Less than Ogive

Table 1.9 : Frequency Distribution more than Method

More than Method	Cf
More than 0	60
More than 10	56
More than 20	51
More than 30	44
More than 40	38
More than 50	28
More than 60	20
More than 70	14
More than 80	9
More than 90	4

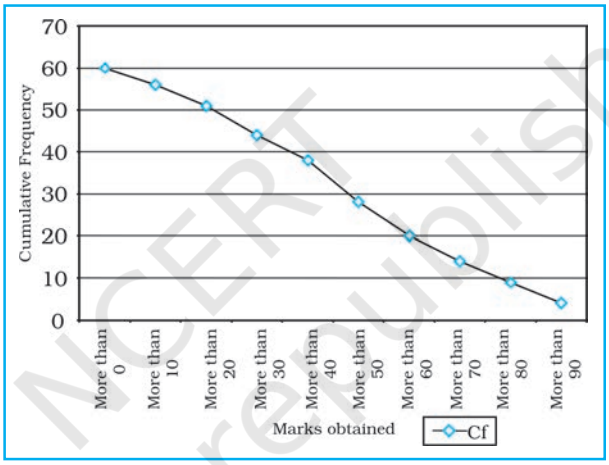


Fig. 1.7 : More than Ogive

Table 1.10 : Less than and more than Ogive

Marks obtained	Less than	More than
0 - 10	4	60
10 - 20	9	56
20 - 30	14	51
30 - 40	21	44
30 - 40	27	38
50 - 60	37	28
60 - 70	45	20
70 - 80	51	14
80 - 90	56	9
90 - 100	60	4

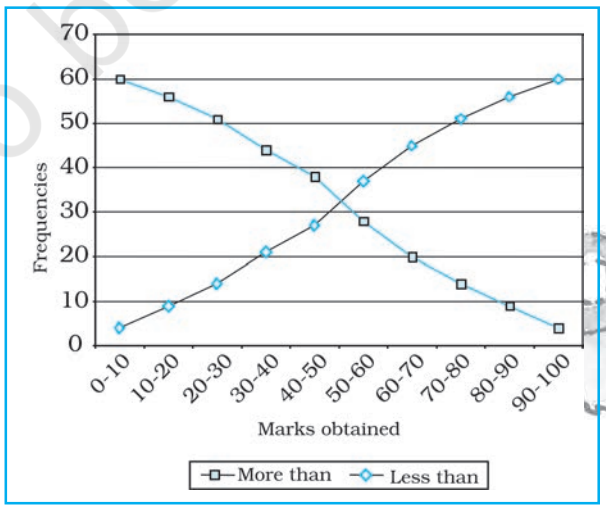


Fig. 1.8 : Less than and more than Ogive

© NCERT not to be republished

Excercises

1. Choose the right answer from the four alternatives given below:
   (i) A number or character which represents measurement is called
   (a) Digit (b) Data (c) Number (d) Character
   (ii) A single datum is a single measurement from the
   (a) Table (b) Frequency (c) Real world (d) Information
   (iii) In a tally mark grouping by four and crossing fifth is called
   (a) Four and Cross Method (b) Tally Marking Method
   (c) Frequency plotting Method (d) Inclusive Method
   (iv) An Ogive is a method in which
   (a) Simple frequency is measured
   (b) Cumulative frequency is measured
   (c) Simple frequency is plotted
   (d) Cumulative frequency is plotted
   (v) If both ends of a group are taken in frequency grouping, it is called
   (a) Exclusive Method (b) Inclusive Method
   (c) Marking Method (d) Statistical Method

2. Answer the following questions in about 30 words:
   (i) Differentiate between data and information.
   (ii) What do you mean by data processing?
   (iii) What is the advantage of foot note in a table?
   (iv) What do you mean by primary sources of data?
   (v) Enumerate five sources of secondary data.

3. Answer the following questions in about 125 words:
   (i) Discuss the national and international agencies where from secondary data may be collected.
   (ii) What is the importance of an index number? Taking an example examine the process of calculating an index number and show the changes.

Activity

1. In a class of 35 students of Geography, following marks were obtained out of 10 marks in unit test – 1, 0, 2, 3, 4, 5, 6, 7, 2, 3, 4, 0, 2, 5, 8, 4, 5, 3, 6, 3, 2, 7, 6, 5, 4, 3, 7, 8, 9, 7, 9, 4, 5, 4, 3. Represent the data in the form of a group frequency distribution.

2. Collect the last test result of Geography of your class and represent the marks in the form of a group frequency distribution.

© NCERT not to be republished

12101CH02

2 Data Processing

You have learnt in previous chapter that organising and presenting data makes them comprehensible. It facilitates data processing. A number of statistical techniques are used to analyse the data e.g.

1. Measures of Central Tendency
2. Measures of Dispersion
3. Measures of Relationship

While measures of central tendency provide the value that is an ideal representative of a set of observations, the measures of dispersion take into account the internal variations of the data, often around a measure of central tendency. The measures of relationship, on the other hand, provide the degree of association between any two or more related phenomena, like rainfall and incidence of flood or fertiliser consumption and yield of crops. In this chapter, you will learn the measures of central tendency.

Measures of Central Tendency

The measurable characteristics such as rainfall, elevation, density of population, levels of educational attainment or age groups vary. If we want to understand them, how would we do ? We may, perhaps, require a single value or number that best represents all the observations. This single value usually lies near the centre of a distribution rather than at either extreme. The statistical techniques used to find out the centre of distributions are referred as **measures of central tendency**. The number denoting the central tendency is the representative figure for the entire data set because it is the point about which items have a tendency to cluster.

Measures of central tendency are also known as statistical averages. There are a number of the measures of central tendency, such as the **mean**, **median** and the **mode**.

Mean

The mean is the value which is derived by summing all the values and dividing it by the number of observations.

© NCERT not to be republished

Median

The median is the value of the rank, which divides the arranged series into two equal numbers. It is independent of the actual value. Arranging the data in ascending or descending order and then finding the value of the middle ranking number is the most significant in calculating the median. In case of the even numbers the average of the two middle ranking values will be the median.

Mode

Mode is the maximum occurrence or frequency at a particular point or value. You may notice that each one of these measures is a different method of determining a single representative number suited to different types of the data sets.

Mean

Mean is the simple arithmetic average of the different values of a variable. For ungrouped and grouped data, the methods for calculating mean are necessarily different. Mean can be calculated by direct or indirect methods, for both grouped and ungrouped data.

Computing Mean from Ungrouped Data

Direct Method

While calculating mean from ungrouped data using the direct method, the values for each observation are added and the total number of occurrences are divided by the sum of all observations. The mean is calculated using the following formula:

$$\bar{X} = \frac{\sum x}{N}$$

Where,

$\bar{X}$ = Mean

$\sum$ = Sum of a series of measures

x = A raw score in a series of measures

$\sum x$ = The sum of all the measures

N = Number of measures

Example 2.1 : Calculate the mean rainfall for Malwa Plateau in Madhya Pradesh from the rainfall of the districts of the region given in *Table 2.1:*

Table 2.1 : Calculation of Mean Rainfall

Districts in Malwa Plateau	*Normal Rainfall in mms*	*Indirect Method*
	***x* Direct Method**	***d= x-800****
Indore	979	179
Dewas	1083	283
Dhar	833	33
Ratlam	896	96
Ujjain	891	91
Mandsaur	825	25
Shajapur	977	177
$\sum x$ and $\sum d$	**6484**	**884**
$\frac{\sum x}{N}$ and $\frac{\sum d}{N}$	**926.29**	**126.29**

* Where 800 is assumed mean.
d is deviation from the assumed mean.

© NCERT not to be republished

The mean for the data given in *Table 2.1* is computed as under:

$$\overline{X} = \frac{\sum x}{N}$$

$$= \frac{6,484}{7}$$

$$= 926.29$$

It could be noted from the computation of the mean that the raw rainfall data have been added directly and the sum is divided by the number of observations i. e., districts. Therefore, it is known as **direct method**.

Indirect Method

For a large number of observations, the indirect method is normally used to compute the mean. It helps in reducing the values of the observations to smaller numbers by subtracting a constant value from them. For example, as shown in *Table 2.1*, the rainfall values lie between 800 and 1100 mm. We can reduce these values by selecting 'assumed mean' and subtracting the chosen number from each value. In the present case, we have taken 800 as assumed mean. Such an operation is known as **coding**. The mean is then worked out from these reduced numbers (Column 3 of *Table 2.1*).

The following formula is used in computing the mean using indirect method:

$$\overline{X} = A + \frac{\sum d}{N}$$

Where,

A = Subtracted constant

$\sum d$ = Sum of the coded scores

N = Number of individual observations in a series

Mean for the data as shown in *Table 2.1* can be computed using the indirect method in the following manner :

$$\overline{X} = 800 + \frac{884}{7}$$

$$= 800 + \frac{884}{7}$$

$$\overline{X} = 926.29 \text{ mm}$$

Note that the mean value comes the same when computed either of the two methods.

Computing Mean from Grouped Data

The mean is also computed for the grouped data using either direct or indirect method.

Direct Method

When scores are grouped into a frequency distribution, the individual values lose their identity. These values are represented by the midpoints of the class

© NCERT not to be republished

intervals in which they are located. While computing the mean from grouped data using direct method, the midpoint of each class interval is multiplied with its corresponding frequency (f); all values of fx (the X are the midpoints) are added to obtain $\sum fx$ that is finally divided by the number of observations i. e., N. Hence, mean is calculated using the following formula :

$$\bar{X} = \frac{\sum fx}{N}$$

Where :

$\bar{X}$ = Mean
f = Frequencies
x = Midpoints of class intervals
N = Number of observations (it may also be defined as $\sum f$)

Example 2.2 : Compute the average wage rate of factory workers using data given in *Table 2.2*:

Table 2.2 : Wage Rate of Factory Workers

Wage Rate (Rs./day)	***Number of workers (f)***
Classes	***f***
50 - 70	10
70 - 90	20
90 - 110	25
110 - 130	35
130 - 150	9

Table 2.3 : Computation of Mean

Classes	***Frequency (f)***	***Mid-points (x)***	***fx***	***d=x-100***	***fd***	***U = (x-100)/ 20***	***fu***
50-70	10	60	600	-40	-400	-2	-20
70-90	20	80	1,600	-20	-400	-1	-20
90-110	**25**	**100**	2,500	**0**	0	0	0
110-130	35	120	4,200	20	700	1	35
130-150	9	140	1,260	40	360	2	18
$\sum fx$ and $\sum fx$	$\sum f$ =99		$\sum fx$ = 10,160		$\sum fd$ = 260		$\sum fu$ = 13

Where N = $\sum f$ = 99

Table 2.3 provides the procedure for calculating the mean for grouped data. In the given frequency distribution, ninety-nine workers have been grouped into five classes of wage rates. The midpoints of these groups are listed in the third column. To find the mean, each midpoint (X) has been multiplied by the frequency (f) and their sum ($\sum fx$) divided by N.

© NCERT not to be republished

The mean may be computed as under using the given formula :

$$\bar{X} = \frac{\sum fx}{N}$$

$$= \frac{10,160}{99}$$

$$= 102.6$$

Indirect Method

The following formula can be used for the indirect method for grouped data. The principles of this formula are similar to that of the indirect method given for ungrouped data. It is expressed as under

$$\bar{x} = A \pm \frac{\sum fd}{N}$$

Where,

A = Midpoint of the assumed mean group
(The assumed mean group in *Table 2.3* is 90 – 110 with 100 as midpoint.)

f = Frequency

d = Deviation from the assumed mean group (A)

N = Sum of cases or $\sum f$

i = Interval width (in this case, it is 20)

From *Table 2.3* the following steps involved in computing mean using the direct method can be deduced :

(i) Mean has been assumed in the group of 90 – 110. It is preferably assumed from the class as near to the middle of the series as possible. This procedure minimises the magnitude of computation. In *Table 2.3*, A (assumed mean) is 100, the midpoint of the class 90 – 110.

(ii) The fifth column (u) lists the deviations of midpoint of each class from the midpoint of the assumed mean group (90 – 110).

(iii) The sixth column shows the multiplied values of each f by its corresponding d to give fd. Then, positive and negative values of fd are added separately and their absolute difference is found ($\sum fd$). Note that the sign attached to $\sum fd$ is replaced in the formula following A, where ± is given.

The mean using indirect method is computed as under :

$$\bar{x} = A \pm \frac{\sum fd}{N}$$

$$= 100 + \frac{260}{99}$$

$$= 100 + 2.6$$

$$= 102.6$$

Note : The Indirect mean method will work for both equal and unequal class intervals.

© NCERT not to be republished

Median

Median is a **positional average**. It may be defined "as the point in a distribution with an equal number of cases on each side of it". The **Median** is expressed using symbol M.

Computing Median for Ungrouped Data

When the scores are ungrouped, these are arranged in ascending or descending order. Median can be found by locating the central observation or value in the arranged series. The central value may be located from either end of the series arranged in ascending or descending order. The following equation is used to compute the median :

$$\text{Value of } \left(\frac{N+1}{2}\right) \text{ th item}$$

Example 2.3: Calculate median height of mountain peaks in parts of the Himalayas using the following:

8,126 m, 8,611m, 7,817 m, 8,172 m, 8,076 m, 8,848 m, 8,598 m.

Computation : Median (M) may be calculated in the following steps :

(i) Arrange the given data in ascending or descending order.

(ii) Apply the formula for locating the central value in the series. Thus :

$$\text{Value of } \left(\frac{N+1}{2}\right) \text{ th item}$$

$$= \left(\frac{7+1}{2}\right) \text{th item}$$

$$= \left(\frac{8}{2}\right) \text{th item}$$

4th item in the arranged series will be the Median.

Arrangement of data in ascending order –

7,817; 8,076; 8,126; 8,172; 8,598; 8,611; 8,848

↓

4th item

Hence,

M = 8,172 m

Computing Median for Grouped Data

When the scores are grouped, we have to find the value of the point where an individual or observation is centrally located in the group. It can be computed using the following formula :

$$M = l + \frac{i}{f}\left(\frac{N}{2} - c\right)$$

© NCERT not to be republished

Where,

M = Median for grouped data
l = Lower limit of the median class
i = Interval
f = Frequency of the median class
N = Total number of frequencies or number of observations
c = Cumulative frequency of the pre-median class.

Example 2.4 : Calculate the median for the following distribution :

class	50-60	60-70	70-80	80-90	90-100	100-110
f	3	7	11	16	8	5

Table 2.4 : Computation of Median

Class	*Frequency (f)*	*Cumulative Frequency (F)*	*Calculation of Median Class*
50-60	3	3	
60-70	7	10	
70-80	11	*21c*	
80-90 (median group)	***16 f***	**37**	$M=\frac{N}{2}$
90-100	8	45	$=\frac{50}{2}$
100-110	5	50	$= 25$
	$\sum f$ **or N= 50**		

The median is computed in the steps given below :

(i) The frequency table is set up as in *Table 2.4*.

(ii) Cumulative frequencies (**F**) are obtained by adding each normal frequency of the successive interval groups, as given in column 3 of *Table 2.4*.

(iii) Median number is obtained by $\frac{N}{2}$ i.e. $\frac{50}{2}$ = **25** in this case, as shown in column 4 of *Table 2.4*.

(iv) Count into the cumulative frequency distribution (**F**) from the top towards bottom until the value next greater than $\frac{N}{2}$ is reached. In this example, $\frac{N}{2}$ is 25, which falls in the Class interval of 40-44 with cumulative frequency of 37, thus the cumulative frequency of the pre-median class is 21 and actual frequency of the median class is 16.

(v) The median is then computed by substituting all the values determined in the step 4 in the following equation :

$$M = l + \frac{i}{f}(m - c)$$

© NCERT not to be republished

$$= 80 + \frac{10}{16}(25 - 21)$$

$$= 80 + \frac{5}{8} \times 4$$

$$= 80 + \frac{5}{2}$$

$$= 80 + 2.5$$

$$M = 82.5$$

Mode

The value that occurs most frequently in a distribution is referred to as **mode**. It is symbolised as **Z** or $\mathbf{M_0}$. Mode is a measure that is less widely used compared to mean and median. There can be more than one type mode in a given data set.

Computing Mode for Ungrouped Data

While computing mode from the given data sets all measures are first arranged in ascending or descending order. It helps in identifying the most frequently occurring measure easily.

Example 2.5 : Calculate mode for the following test scores in geography for ten students :

61, 10, 88, 37, 61, 72, 55, 61, 46, 22

Computation : To find the mode the measures are arranged in ascending order as given below:

10, 22, 37, 46, 55, **61, 61, 61**, 72, 88.

The measure 61 occurring three times in the series is the **mode** in the given dataset. As no other number is in the similar way in the dataset, it possesses the property of being **unimodal**.

Example 2.6 : Calculate the mode using a different sample of ten other students, who scored:

82, 11, 57, 82, 08, 11, 82, 95, 41, 11.

Computation : Arrange the given measures in an ascending order as shown below :

08, 11, 11, 11, 41, 57, 82, 82, 82, 95

It can easily be observed that measures of 11 and 82 both are occurring three times in the distribution. The dataset, therefore, is **bimodal** in appearance. If three values have equal and highest frequency, the series is **trimodal**. Similarly, a recurrence of many measures in a series makes it **multimodal**. However, when there is no measure being repeated in a series it is designated as **without mode**.

Comparison of Mean, Median and Mode

The three measures of the **central tendency** could easily be compared with the help of normal distribution curve. The normal curve refers to a frequency distribution in which the graph of scores often called a bell-shaped curve. Many

© NCERT not to be republished

human traits such as intelligence, personality scores and student achievements have normal distributions. The bell-shaped curve looks the way it does, as it is symmetrical. In other words, most of the observations lie on and around the middle value. As one approaches the extreme values, the number of observations reduces in a symmetrical manner. A normal curve can have high or low data variability. An example of a normal distribution curve is given in *Fig. 2.1*.

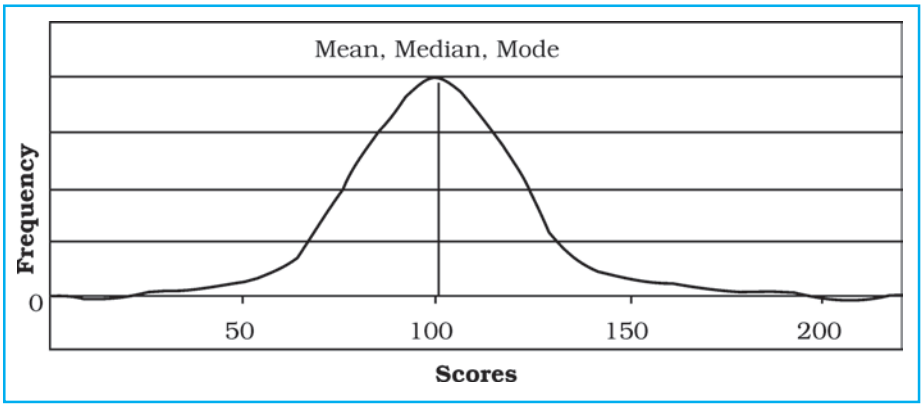


Fig. 2.1 : *Normal Distribution Curve*

The normal distribution has an important characteristic. **The mean, median and mode are the same score** (a score of 100 in *Fig. 2.1*) because a normal distribution is symmetrical. The score with the highest frequency occurs in the middle of the distribution and exactly half of the scores occur above the middle and half of the scores occur below. Most of the scores occur around the middle of the distribution or the mean. Very high and very low scores do not occur frequently and are, therefore, considered rare.

If the data are skewed or distorted in some way, the mean, median and mode will not coincide and the effect of the skewed data needs to be considered (*Fig. 2.2 and 2.3*).

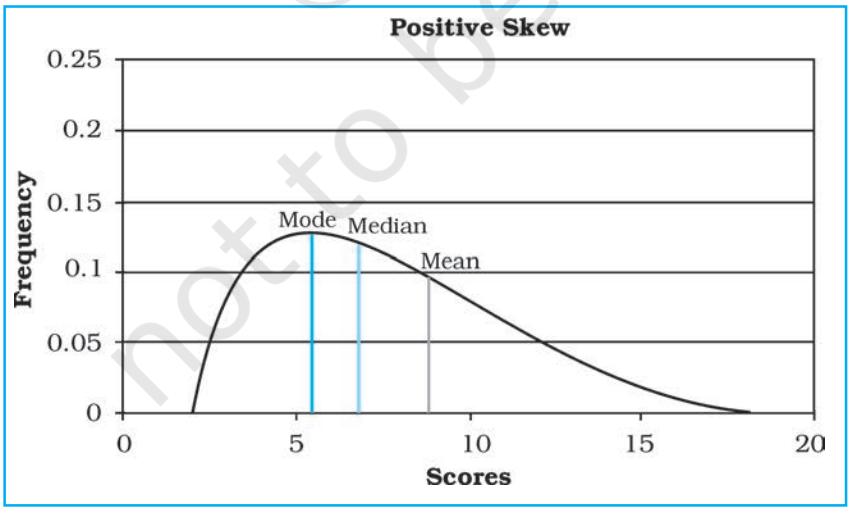


Fig. 2.2 : *Positive Skew*

© NCERT not to be republished

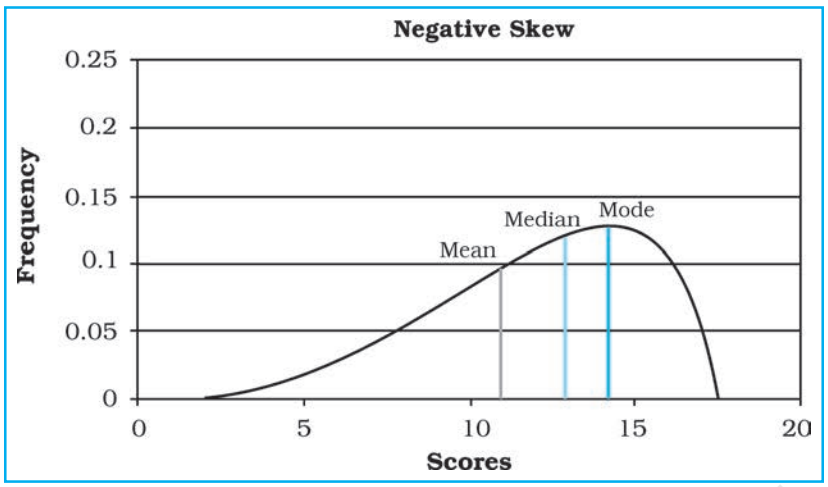


Fig. 2.3 : Negative Skew

Excercises

1. Choose the correct answer from the four alternatives given below:
   (i) The measure of central tendency that does not get affected by extreme values:
   (a) Mean (b) Mean and Mode
   (c) Mode (d) Median
   (ii) The measure of central tendency always coinciding with the hump of any distribution is:
   (a) Median (b) Median and Mode
   (c) Mean (d) Mode
2. Answer the following questions in about 30 words:
   (i) Define the mean.
   (ii) What are the advantages of using mode ?
3. Answer the following questions in about 125 words:
   (i) Explain relative positions of mean, median and mode in a normal distribution and skewed distribution with the help of diagrams.
   (ii) Comment on the applicability of mean, median and mode (*hint: from their merits and demerits).*

Activity

1. Take an imaginary example applicable to geographical analysis and explain direct and indirect methods of calculating mean from ungrouped data.

© NCERT not to be republished

12101CH03

3 Graphical Representation of Data

You must have seen graphs, diagrams and maps showing different types of data. For example, the thematic maps shown in Chapter 1 of book for Class XI entitled *Practical Work in Geography, Part-I (NCERT, 2006)* depict relief and slope, climatic conditions, distribution of rocks and minerals, soils, population, industries, general land use and cropping pattern in the Nagpur district, Maharashtra. These maps have been drawn using large volume of related data collected, compiled and processed. Have you ever thought what would have happened if the same information would have been either in tabular form or in a descriptive transcript? Perhaps, it would not have been possible from such a medium of communication to draw visual impressions which we get through these maps. Besides, it would also have been a time consuming task to draw inferences about whatever is being presented in non–graphical form. Hence, the graphs, diagrams and maps enhance our capabilities to make meaningful comparisons between the phenomena represented, save our time and present a simplified view of the characteristics represented. In the present chapter, we will discuss methods of constructing different types of graphs, diagrams and maps.

Representation of Data

The data describe the properties of the phenomena they represent. They are collected from a variety of sources (Chapter 1). The geographers, economists, resource scientists and the decision makers use a lot of data these days. Besides the tabular form, the data may also be presented in some graphic or diagrammatic form. The transformation of data through visual methods like graphs, diagrams, maps and charts is called representation of data. Such a form of the presentation of data makes it easy to understand the patterns of population growth, distribution and the density, sex ratio, age–sex composition, occupational structure, etc. within a geographical territory. There is a Chinese proverb that *'a picture is equivalent to thousands of words'*. Hence, the graphic method of the representation of data enhances our understanding, and makes the comparisons easy. Besides, such methods create an imprint on mind for a longer time.

© NCERT not to be republished

General Rules for Drawing Graphs, Diagrams and Maps

1. Selection of a Suitable Method

Data represent various themes such as temperature, rainfall, growth and distribution of the population, production, distribution and trade of different commodities, etc. These characteristics of the data need to be suitably represented by an appropriate graphical method. For example, data related to the temperature or growth of population between different periods in time and for different countries/states may best be represented using line graphs. Similarly, bar diagrams are suited best for showing rainfall or the production of commodities. The population distribution, both human and livestock, or the distribution of the crop producing areas may suitably be represented on dot maps and the population density using choropleth maps.

2. Selection of Suitable Scale

The scale is used as measure of the data for representation over diagrams and maps. Hence, the selection of suitable scale for the given data sets should be carefully made and must take into consideration entire data that is to be represented. The scale should neither be too large nor too small.

3. Design

We know that the design is an important cartographic task (Refer 'Essentials of Map Making' as discussed in Chapter 1 of the *Practical Work in Geography, Part-I (NCERT, 2006),* a textbook of Class XI). The following components of the cartographic designs are important. Hence, these should be carefully shown on the final diagram/map.

Title

The title of the diagram/map indicates the name of the area, reference year of the data used and the caption of the diagram. These components are represented using letters and numbers of different font sizes and thickness. Besides, their placing also matters. Normally, title, subtitle and the corresponding year are shown in the centre at the top of the map/diagram.

Legend

A legend or index is an important component of any diagram/map. It explains the colours, shades, symbols and signs used in the map and diagram. It should also be carefully drawn and must correspond to the contents of the map/diagram. Besides, it also needs to be properly positioned. Normally, a legend is shown either at the lower left or lower right side of the map sheet.

Direction

The maps, being a representation of the part of the earth's surface, need be oriented to the directions. Hence, the direction symbol, i. e. North, should also be drawn and properly placed on the final map.

Construction of Diagrams

The data possess measurable characteristics such as length, width and volume. The diagrams and the maps that are drawn to represent these data related characteristics may be grouped into the following types:

© NCERT not to be republished

(i) One-dimensional diagrams, such as line graph, poly graph, bar diagram, histogram, age, sex, pyramid, etc.;

(ii) Two-dimensional diagram, such as pie diagram and rectangular diagram;

(iii) Three-dimensional diagrams, such as cube and spherical diagrams.

It would not be possible to discuss the methods of construction of these many types of diagrams and maps primarily due to the time constraint. We will, therefore, describe the most commonly drawn diagrams and maps and the way they are constructed. These are :

- Line graphs
- Pie diagram
- Bar diagrams
- Wind rose and star diagram
- Flow Charts

Line Graph

The line graphs are usually drawn to represent the time series data related to the temperature, rainfall, population growth, birth rates and the death rates. *Table 3.1* provides the data used for the construction of Fig 3.2.

Construction of a Line Graph

(a) Simplify the data by converting it into round numbers, such as the growth rate of population as shown in *Table 3.1* for the years 1961 and 1981 may be rounded to 2.0 and 2.2, respectively.

(b) Draw X and Y-axis. Mark the time series variables (years/months) on the X axis and the data quantity/value to be plotted (growth of population in per cent or the temperature in ˚C) on Y axis.

(c) Choose an appropriate scale and label it on Y-axis. If the data involve a negative figure, then the selected scale should also show it as shown in *Fig. 3.1*.

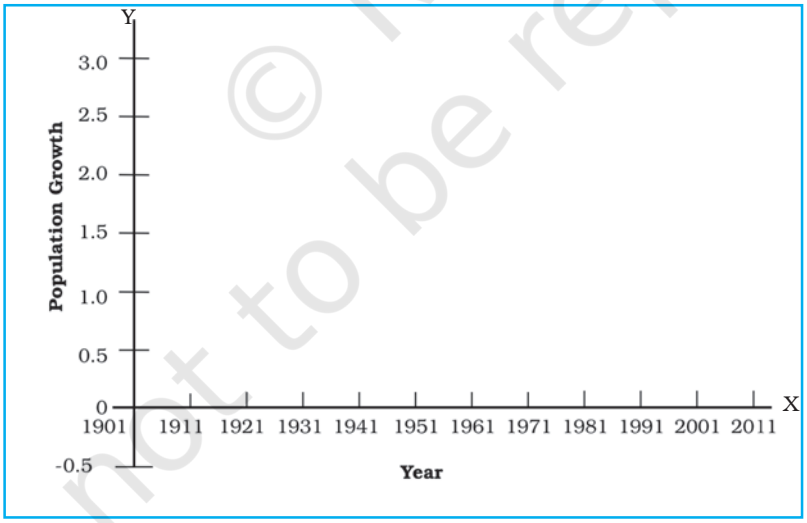


Fig. 3.1 : *Construction of a Line Graph*

(d) Plot the data to depict year/month-wise values according to the selected scale on Y-axis, mark the location of the plotted values by a dot and join these dots by a free hand drawn line.

© NCERT not to be republished

Example 3.1 : Construct a line graph to represent the data as given in *Table 3.1*:

Table 3.1 : Growth rate of Population in India – 1901 to 2011

Year	Growth rate in percentage
1901	-
1911	0.56
1921	-0.30
1931	1.04
1941	1.33
1951	1.25
1961	1.96
1971	2.20
1981	2.22
1991	2.14
2001	1.93
2011	1.79

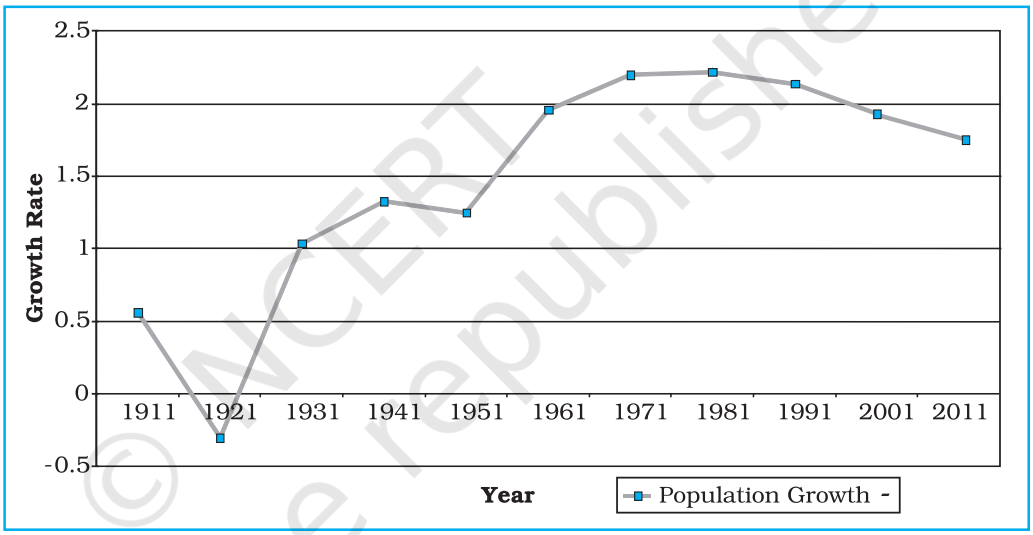


Fig. 3.2 : *Annual Growth of Population in India 1901-2011*

Activity

Find out the reasons for sudden change in population between 1911 and 1921as shown in *Fig. 3.2*.

Polygraph

Polygraph is a line graph in which two or more than two variables are shown by an equal number of lines for an immediate comparison, such as the growth rate of different crops like rice, wheat, pulses or the birth rates, death rates and life expectancy or sex ratio in different states or countries. A different line pattern such as straight line (____), broken line (- - -), dotted line (……) or a combination of dotted and broken line (-.-.-) or line of different colours may be used to indicate the value of different variables (*Fig 3.3*).

© NCERT not to be republished

Example 3.2 : Construct a polygraph to compare the growth of sex-ratio in different states as given in the *Table 3.2* :

Table 3.2 : Sex-Ratio (Female per 1000 male) of Selected Sates – 1961-2011

States/UT	*1961*	*1971*	*1981*	*1991*	*2001*	*2011*
Delhi	785	801	808	827	821	866
Haryana	868	867	870	860	846	877
Uttar Pradesh	907	876	882	876	898	908

Source : Census, 2011

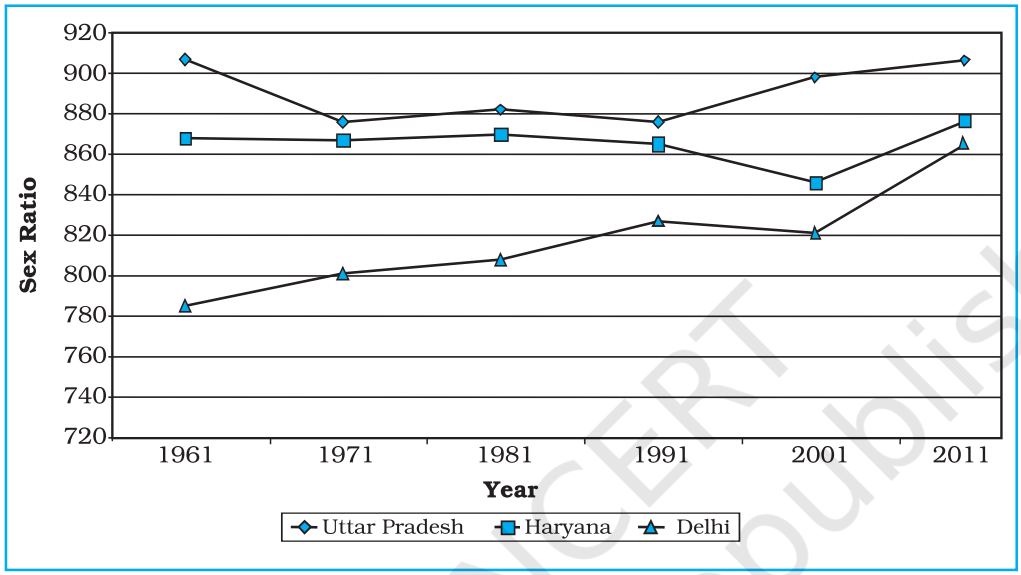


Fig. 3.3 : *Sex-Ratio of Selected States 1961-2011*

Bar Diagram

The bar diagrams are drawn through columns of equal width. It is also called a columnar diagram. Following rules should be observed while constructing a bar diagram:

(a) The width of all the bars or columns should be similar.

(b) All the bars should be placed on equal intervals/distance.

(c) Bars may be shaded with colours or patterns to make them distinct and attractive.

The simple, compound or polybar diagram may be constructed to suit the data characteristics.

Simple Bar Diagram

A simple bar diagram is constructed for an immediate comparison. It is advisable to arrange the given data set in an ascending or descending order and plot the data variables accordingly. However, time series data are represented according to the sequencing of the time period.

Example 3.3 : Construct a simple bar diagram to represent the rainfall data of Thiruvananthapuram as given in *Table 3.3* :

© NCERT not to be republished

Table 3.3 : Average Monthly Rainfall of Thiruvananthapuram

Months	J	F	M	A	M	J	J	A	S	O	N	D
Rainfall in cm	2.3	2.1	3.7	10.6	20.8	35.6	22.3	14.6	13.8	27.3	20.6	7.5

Construction

Draw X and Y-axes on a graph paper. Take an interval of 5 cm and mark it on Y-axis to plot rainfall data in cm. Divide X-axis into 12 equal parts to represent 12 months. The actual rainfall values for each month will be plotted according to the selected scale as shown in *Fig. 3.4*.

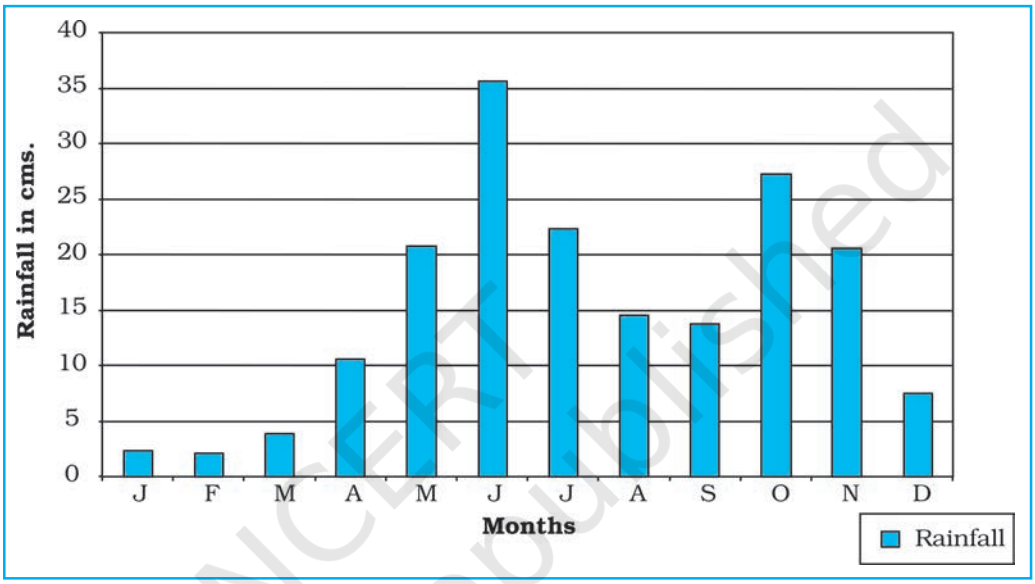


Fig. 3.4 : *Average Monthly Rainfall of Thiruvananthapuram*

Line and Bar Graph

The line and bar graphs as drawn separately may also be combined to depict the data related to some of the closely associated characteristics such as the climatic data of mean monthly temperatures and rainfall. In doing so, a single diagram is drawn in which months are represented on X-axis while temperature and rainfall data are shown on Y-axis at both sides of the diagram.

Example 3.4 : Construct a line graph and bar diagram to represent the average monthly rainfall and temperature data of Delhi as given in *Table 3.4* :

Table 3.4 : Average monthly Temperature and Rainfall in Delhi

Months	***Temp. in °C***	***Rainfall in cm.***
Jan.	14.4	2.5
Feb.	16.7	1.5
Mar.	23.30	1.3
Apr.	30.0	1.0
May	33.3	1.8
June	33.3	7.4
Jul.	30.0	19.3
Aug.	29.4	17.8
Sep.	28.9	11.9
Oct.	25.6	1.3
Nov.	19.4	0.2
Dec.	15.6	1.0

© NCERT not to be republished

Construction

(a) Draw X and Y-axes of a suitable length and divide X-axis into 12 parts to show months in a year.

(b) Select a suitable scale with equal intervals of 5° C or 10° C for temperature data on the Y-axis and label it at its right side.

(c) Similarly, select a suitable scale with equal intervals of 5 cm or 10 cm for rainfall data on the Y-axis and label at its left side.

(d) Plot temperature data using line graph and the rainfall by bar diagram as shown in *Fig. 3.5*.

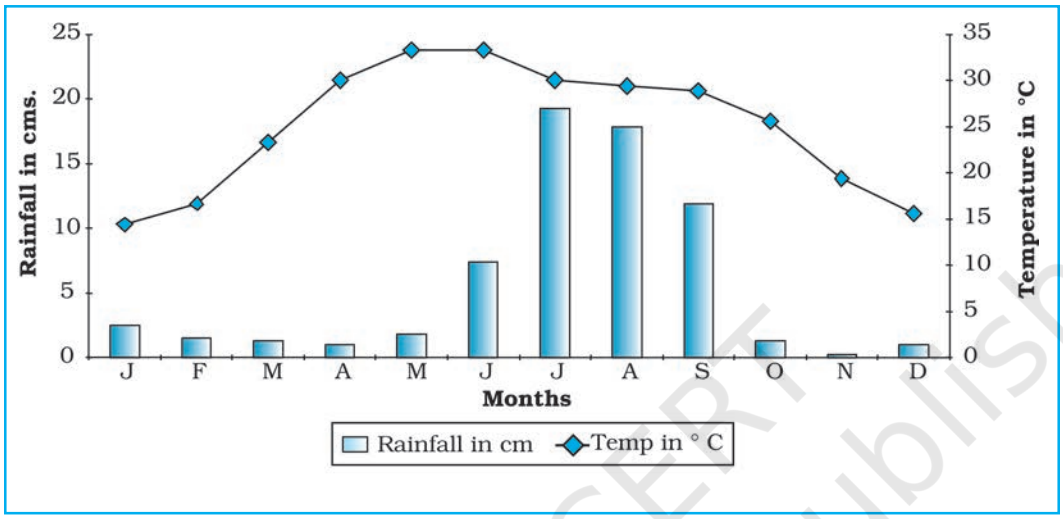


Fig. 3.5 : *Temperature and Rainfall in Delhi*

Multiple Bar Diagram

Multiple bar diagrams are constructed to represent two or more than two variables for the purpose of comparison. For example, a multiple bar diagram may be constructed to show proportion of males and females in the total, rural and urban population or the share of canal, tube well and well irrigation in the total irrigated area in different states.

Example 3.5 : Construct a suitable bar diagram to show decadal literacy rate in India during 1951–2011 as given in *Table 3.5* :

Table 3.5 : Literacy Rate in India, 1951–2011 (in %)

Year	Literacy Rate		
	Total population	Male	Female
1951	18.33	27.16	8.86
1961	28.3	40.4	15.35
1971	34.45	45.96	21.97
1981	43.57	56.38	29.76
1991	52.21	64.13	39.29
2001	64.84	75.85	54.16
2011	73.0	80.9	64.6

Source : Census, 2011

Construction

(a) Multiple bar diagram may be chosen to represent the above data.

(b) Mark time series data on X-axis and literacy rates on Y-axis as per the selected scale.

© NCERT not to be republished

(c) Plot the per cent of total population, male and female in closed columns (*Fig 3.6*).

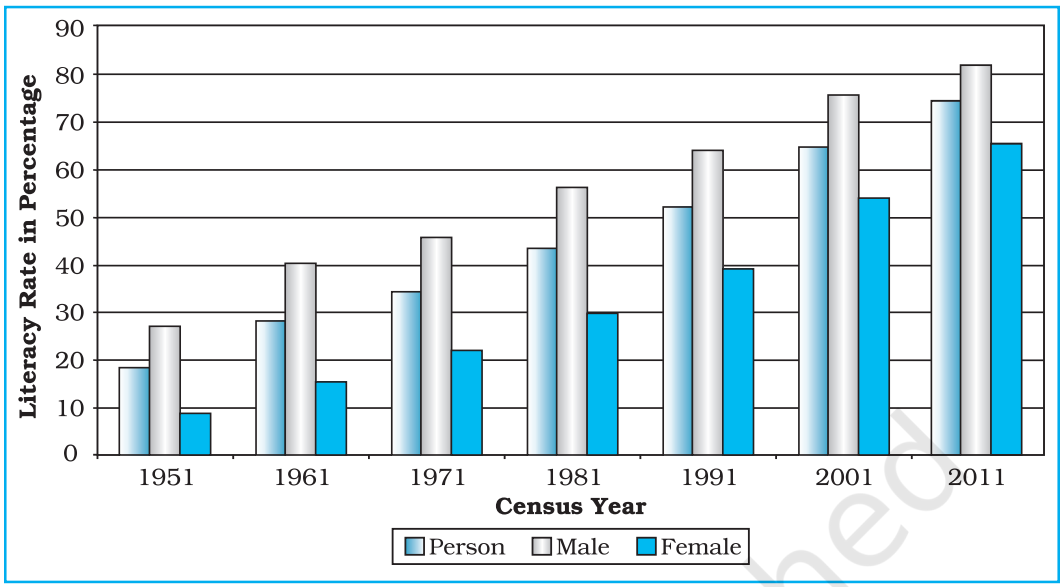


Fig. 3.6 : *Literacy Rate in India, 1951-2011*

Compound Bar Diagram

When different components are grouped in one set of variable or different variables of one component are put together, their representation is made by a compound bar diagram. In this method, different variables are shown in a single bar with different rectangles.

Example 3.6 : Construct a compound bar diagram to depict the data as shown in *Table 3.6* :

Table 3.6 : Gross Generation of Electricity in India (in Billion KWh)

Year	Thermal	Hydro	Nuclear	Total
2008-09	616.2	110.1	14.9	741.2
2009-10	677.1	104.1	18.6	799.8
2010-11	704.3	114.2	26.3	844.8

Source: Economic Survey, 2011-12

Construction

(a) Arrange the data in ascending or descending order.

(b) A single bar will depict the gross electricity generation in the given year and the generation of thermal, hydro and nuclear electricity be shown by dividing the total length of the bar as shown in *Fig 3.7*.

© NCERT not to be republished

Pie Diagram

Pie diagram is another graphical method of the representation of data. It is drawn to depict the total value of the given attribute using a circle. Dividing the circle into corresponding degrees of angle then represent the sub-sets of the data. Hence, it is also called **Divided Circle Diagram**.

The angle of each variable is calculated using the following formulae.

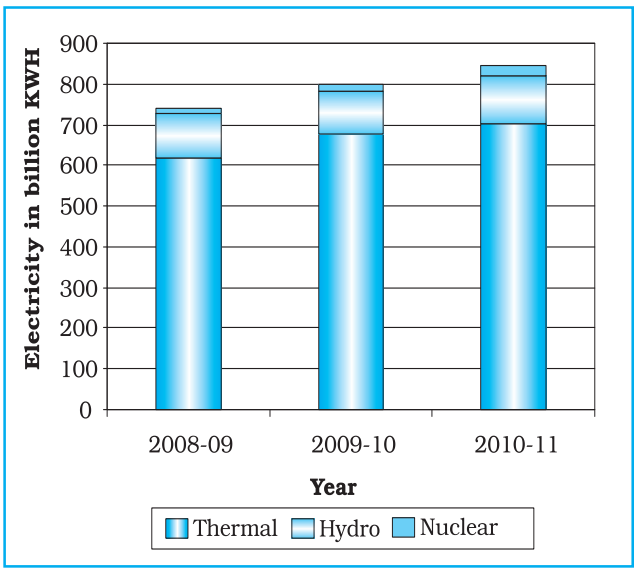


Fig. 3.7 : Gross Electricity Generation in India

$$\frac{\text{Value of given State/Region} \times 360}{\text{Total Value of All States/Regions}}$$

If data are given in percentage form, the angles are calculated using the given formulae.

$$\frac{\text{Percentage of x} \times 360}{100}$$

For example, a pie diagram may be drawn to show the total population of India along with the proportion of the rural and urban population. In this case, the circle of an appropriate radius is drawn to represent the total population and its sub-divisions into rural and urban population are shown by corresponding degrees of angle.

Example 3.7: Represent the data as given in *Table 3.7 (a)* with a suitable diagram.

Table 3.7 (a) : India's Export to Major Regions of the World in 2010–11

Unit/Region	*% of Indian Export*
Europe	20.2
Africa	6.5
America	14.8
Asia and ASEAN	56.2
Others	2.3
Total	100

Source : Economic Survey 2011-12

Calculation of Angles

(a) Arrange the data on percentages of Indian exports in an ascending order.

(b) Calculate the degrees of angles for showing the given values of India's export to major regions/ countries of the world, *Table 3.7* (b). It could be done by multiplying percentage with a constant of 3.6 as derived by dividing the total number of degrees in a circle by 100, i. e. 360/100.

© NCERT not to be republished

(c) Plot the data by dividing the circle into the required number of divisions to show the share of India's export to different regions/countries (*Fig. 3.8*).

Table 3.7 (b) : India's Export to Major Regions of the World in 2010-11

Countries	%	Calculation	Degree
Europe	20.2	20.2 × 3.6 = 72.72	73°
Africa	6.5	6.5 × 3.6 = 23.4	23°
America	14.8	14.8 × 3.6 = 53.28	53°
Asia and ASEAN	56.2	56.2 × 3.6 = 202.32	203°
Others	2.3	2.3 × 3.6 = 8.28	8°
Total	**100**		**360°**

Construction

(a) Select a suitable radius for the circle to be drawn. A radius of 3, 4 or 5 cm may be chosen for the given data set.

(b) Draw a line from the centre of the circle to the arc as a radius.

(c) Measure the angles from the arc of the circle for each category of vehicles in an ascending order clock-wise, starting with smaller angle.

(d) Complete the diagram by adding the title, sub-title, and the legend. The legend mark be chosen for each variable/category and highlighted by distinct shades/colours.

Precautions

(a) The circle should neither be too big to fit in the space nor too small to be illegible.

(b) Starting with bigger angle will lead to accumulation of error leading to the plot of the smaller angle difficult.

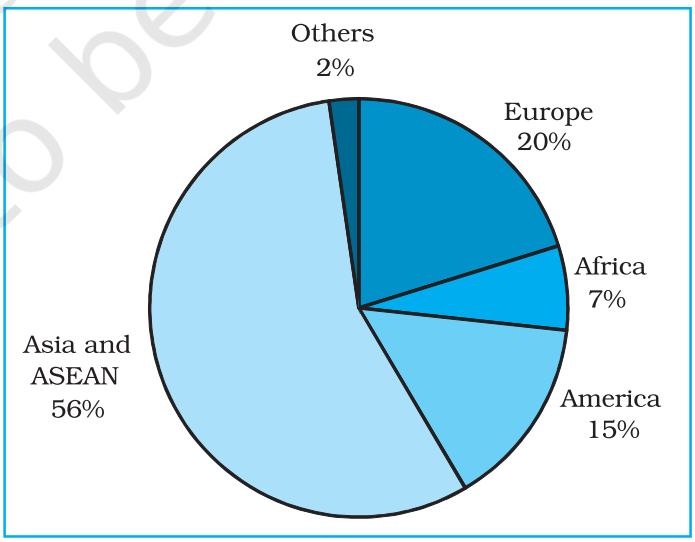


Fig. 3.8 : *Direction of Indian Exports 2010–11*

© NCERT not to be republished

Flow Maps/Chart

Flow chart is a combination of graph and map. It is drawn to show the flow of commodities or people between the places of origin and destination. It is also called **Dynamic Map**. Transport map, which shows the number of passengers, vehicles, etc., is the best example of a flow chart. These charts are drawn using lines of proportional width. Many government agencies prepare flow maps to show density of the means of transportation on different routes. The flow maps/ charts are generally drawn to represent two the types of data as given below:

1. The number and frequency of the vehicles as per the direction of their movement
2. The number of the passengers and/or the quantity of goods transported.

Requirements for the Preparation of a Flow Map

(a) A route map depicting the desired transport routes along with the connecting stations.

(b) The data pertaining to the flow of goods, services, number of vehicles, etc., along with the point of origin and destination of the movements.

(c) The selection of a scale through which the data related to the quantity of passengers and goods or the number of vehicles is to be represented.

Example 3.8 : Construct a flow map to represent the number of trains running in Delhi and the adjoining areas as given in the *Table 3.8*.

Table 3.8 : No. of trains of selected routes of Delhi and adjoining areas

S. No.	Railway Routes	No. of Trains
1.	Old Delhi – New Delhi	50
2.	New Delhi-Nizamuddin	40
3.	Nizamuddin-Badarpur	30
4.	Nizamuddin-Sarojini Nagar	12
5.	Sarojini Nagar – Pusa Road	8
6.	Old Delhi – Sadar Bazar	32
7.	Udyog Nagar-Tikri Kalan	6
8.	Pusa Road – Pehladpur	15
9.	Sahibabad-Mohan Nagar	18
10.	Old Delhi – Silampur	33
11.	Silampur – Nand Nagari	12
12.	Silampur-Mohan Nagar	21
13.	Old Delhi-Shalimar Bagh	16
14.	Sadar Bazar-Udyog Nagar	18
15.	Old Delhi – Pusa Road	22
16.	Pehladpur – Palam Vihar	12

Construction

(a) Take an outline map of Delhi and adjoining areas, in which railway line and the nodal stations are depicted (*Fig.3.9*).

(b) Select a scale to represent the number of trains. Here, the maximum number is 50 and the minimum is 6. If we select a scale of 1cm = 50 trains, the maximum and minimum numbers will be represented by a strip of 10 mm and 1.2 mm thick lines, respectively, on the map.

(c) Plot the thickness of each strip of route between the given rail route (*Fig. 3.10*).

© NCERT not to be republished

(d) Draw a terraced scale as legend and choose distinct sign or symbol to show the nodal points (stations) within the strip.

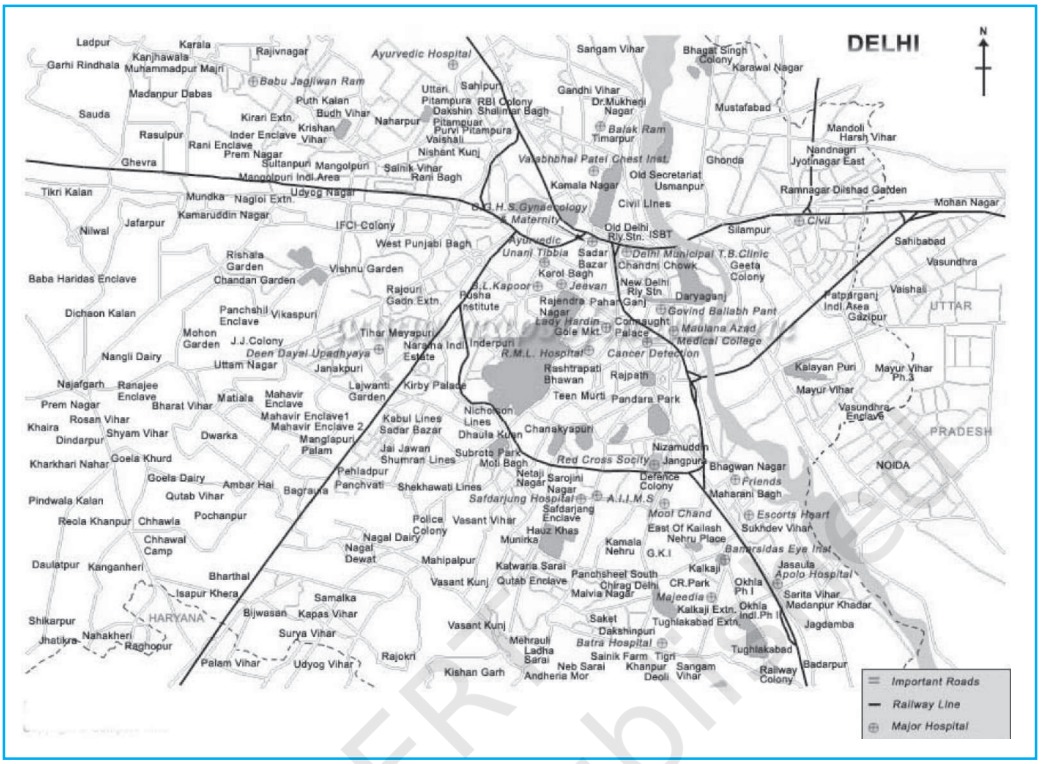


Fig. 3.9 : Map of Delhi

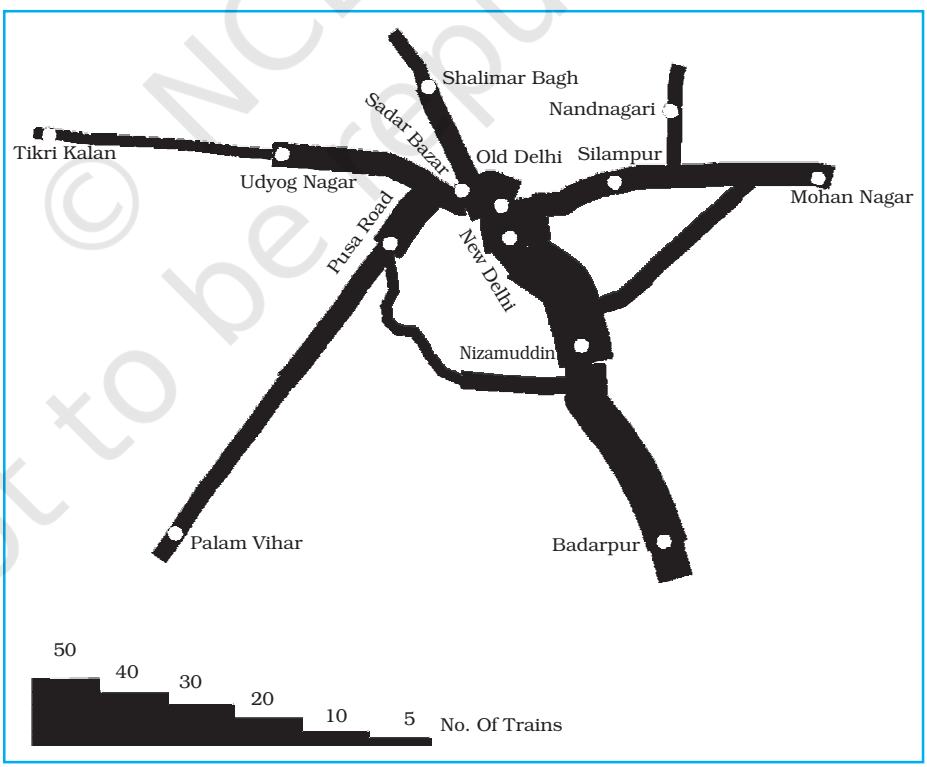


Fig. 3.10 : Traffic (Railway) Flow Map of Delhi

© NCERT not to be republished

Example 3.9 : Construct a water flow map of Ganga Basin as shown in *Fig. 3.11.*

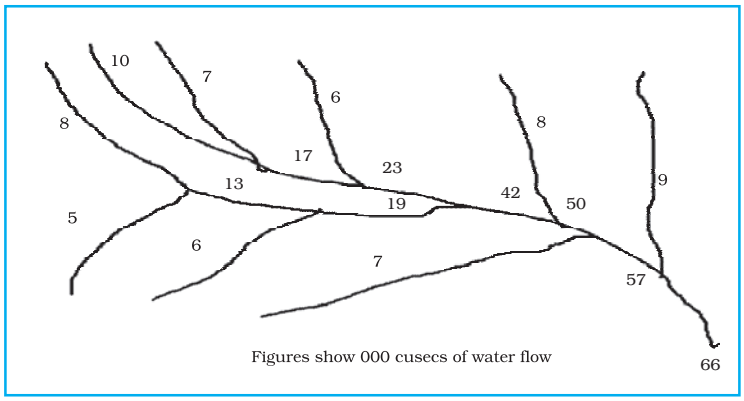


Fig. 3.11 : *Ganga Basin*

Construction

(a) Take a scale as a strip of 1cm width = 50,000 cusecs of water.

(b) Make the diagram as shown in *Fig. 3.12.*

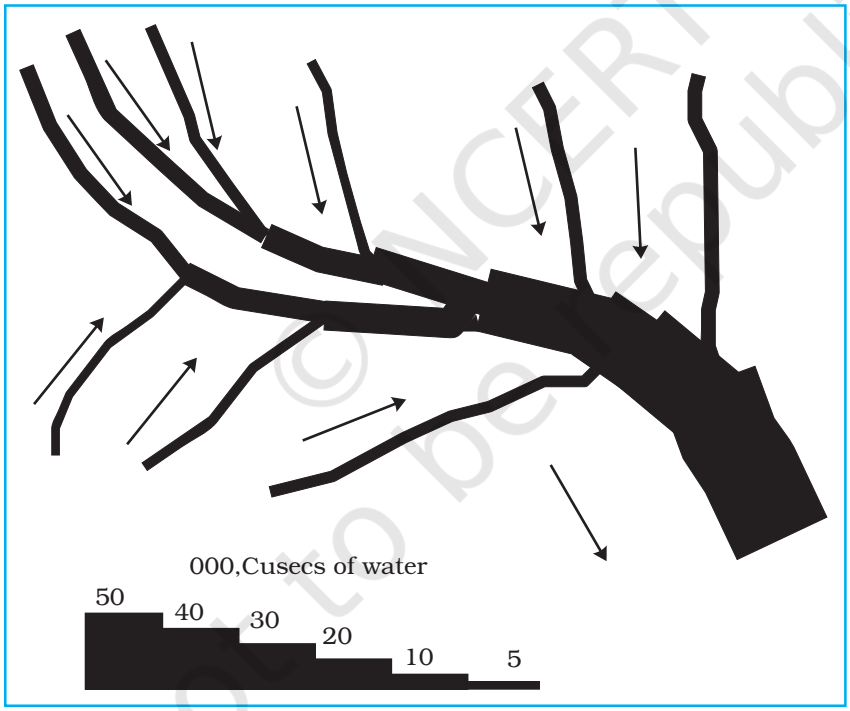


Fig. 3.12 : *Construction of a Flow Map*

Thematic Maps

Graphs and diagrams serve a useful purpose in providing a comparison between the internal variations within the data of different characteristics represented. However, the use of graphs and diagrams, at times, fails to produce a regional perspective. Hence, variety of maps may also be drawn to understand the patterns

© NCERT not to be republished

of the regional distributions or the characteristics of variations over space. These maps are also known as the **distribution maps**.

Requirements for Making a Thematic Map

(a) State/District level data about the selected theme.

(b) Outline map of the study area alongwith administrative boundaries.

(c) Physical map of the region. For example, physiographic map for population distribution and relief and drainage map for constructing transportation map.

Rules for Making Thematic Maps

(i) The drawing of the thematic maps must be carefully planned. The final map should properly reflect the following components:

a. Name of the area
b. Title of the subject-matter
c. Source of the data and year
d. Indication of symbols, signs, colours, shades, etc.
e. Scale

(ii) The selection of a suitable method to be used for thematic mapping.

Classification of Thematic Maps based on Method of Construction

The thematic maps are, generally, classified into quantitative and non-quantitative maps. The quantitative maps are drawn to show the variations within the data. For example, maps depicting areas receiving more than 200 cm, 100 to 200 cm, 50 to 100 cm and less than 50 cm of rainfall are referred as quantitative maps. These maps are also called statistical maps. The non-quantitative maps, on the other hand, depict the non-measurable characteristics in the distribution of given information, such as a map showing high and low rainfall-receiving areas. These maps are also called qualitative maps. It would not be possible to discuss drawing these different types of thematic maps under the constraint of time. We will, therefore, confine to discuss the methods of the construction of the following types of quantitative maps :

(a) Dot maps

(b) Choropleth maps

(c) Isopleth maps

Dot Maps

The dot maps are drawn to show the distribution of phenomena such as population, cattle, types of crops, etc. The dots of same size as per the chosen scale are marked over the given administrative units to highlight the patterns of distributions.

Requirement

(a) An administrative map of the given area showing state/district/block boundaries.

© NCERT not to be republished

(b) Statistical data on selected theme for the chosen administrative units, i.e., total population, cattle, etc.
(c) Selection of a scale to determine the value of a dot.
(d) Physiographic map of the region, especially relief and drainage maps.

Precaution

(a) The lines, demarcating the boundaries of various administrative units, should not be very thick and bold.
(b) All dots should be of same size.

Example 3.10 : Construct a dot map to represent population data of 2001 as given in *Table 3.9*.

Table 3.9 : Population of India, 2001

Sl. No.	States/Union Territories	Total Population	No. of dots
1.	Jammu & Kashmir	10,069,917	100
2.	Himachal Pradesh	6,077,248	60
3.	Punjab	24,289,296	243
4.	Uttarakhand	8,479,562	85
5.	Haryana	21,082,989	211
6.	Delhi	13,782,976	138
7.	Rajasthan	56,473,122	565
8.	Uttar Pradesh	166,052,859	1,660
9.	Bihar	82,878,796	829
10.	Sikkim	540,493	5
11.	Arunachal Pradesh	1,091,117	11
12.	Nagaland	1,988,636	20
13.	Manipur	2,388,634	24
14.	Mizoram	891,058	89
15.	Tripura	3,191,168	32
16.	Meghalaya	2,306,069	23
17.	Assam	26,638,407	266
18.	West Bengal	80,221,171	802
19.	Jharkhand	26,909,428	269
20.	Odisha	36,706,920	367
21.	Chhattisgarh	20,795,956	208
22.	Madhya Pradesh	60,385,118	604
23.	Gujarat	50,596,992	506
24.	Maharashtra	96,752,247	968
25.	Andhra Pradesh	75,727,541	757
26.	Karnataka	52,733,958	527
27.	Goa	1,343,998	13
28.	Kerala	31,838,619	318
29.	Tamil Nadu	62,110,839	621

© NCERT not to be republished

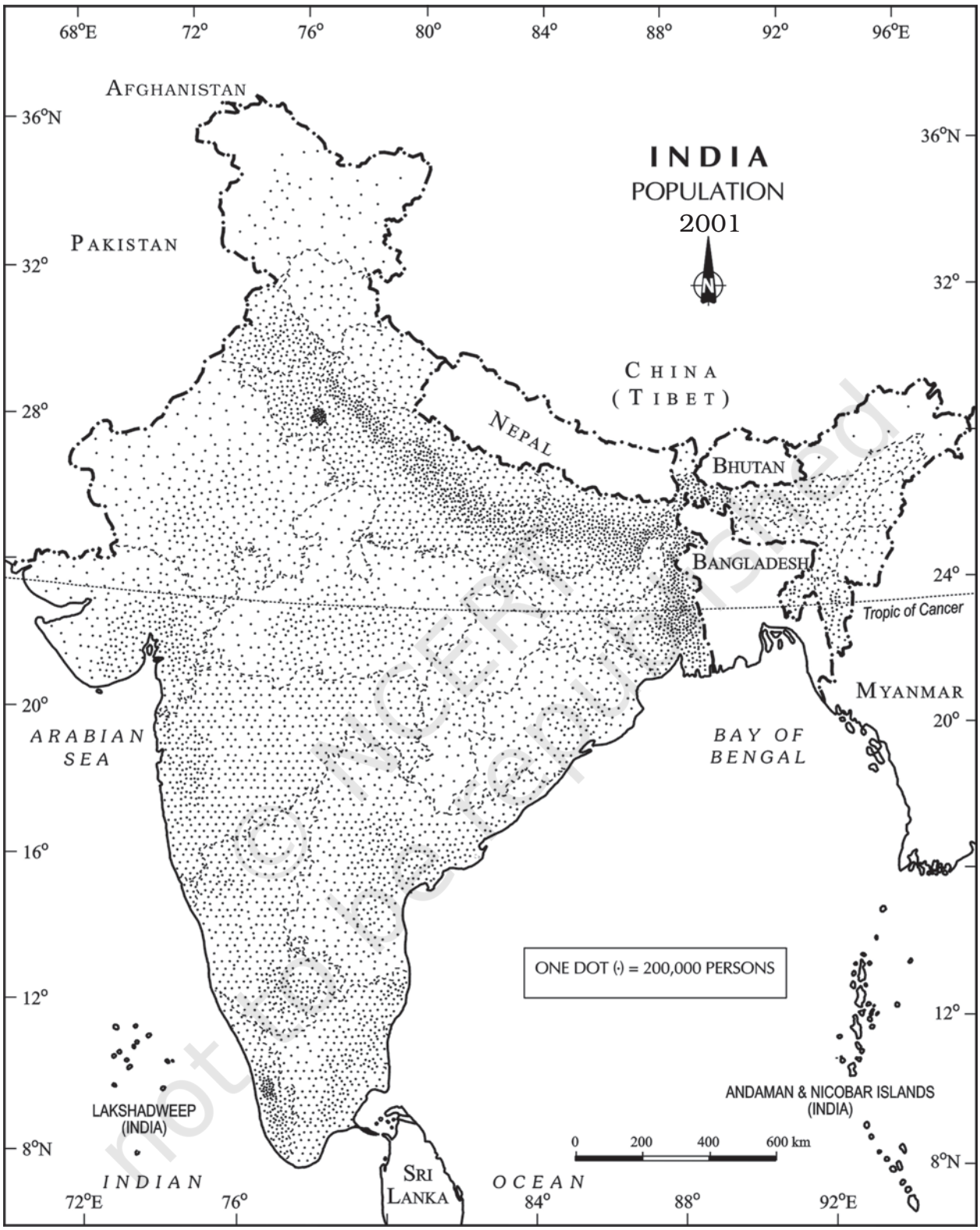


Fig. 3.13 : Population of India, 2001

© NCERT not to be republished

Construction

(a) Select the size and value of a dot.

(b) Determine the number of dots in each state using the given scale. For example, number of dots in Maharashtra will be 9,67,52,247/100,000 = 967.52. It may be rounded to 968, as the fraction is more than 0.5.

(c) Place the dots in each state as per the determined number in all states.

(d) Consult the physiographic/relief map of India to identify mountainous, desert, and/or snow covered areas and mark lesser number of dots in such areas.

Choropleth Map

The choropleth maps are also drawn to depict the data characteristics as they are related to the administrative units. These maps are used to represent the density of population, literacy/growth rates, sex ratio, etc.

Requirement for drawing Choropleth Map

(a) A map of the area depicting different administrative units.

(b) Appropriate statistical data according to administrative units.

Steps to be followed

(a) Arrange the data in ascending or descending order.

(b) Group the data into 5 categories to represent very high, high, medium, low and very low concentrations.

(c) The interval between the categories may be identified on the following formulae i.e., Range/5 and Range = maximum value – minimum value.

(d) Patterns, shades or colour to be used to depict the chosen categories should be marked in an increasing or decreasing order.

Example 3.11: Construct a Choropleth map to represent the literacy rates in India in 2001 as given in *Table 3.10*.

Construction

(a) Arrange the data in ascending order as shown above.

(b) Identify the range within the data. In the present case, the states recording the lowest and highest literacy rates are Bihar (47%) and Kerala (90.9%), respectively. Hence, the range would be 91.0 – 47.0 = 44.0

(c) Divide the range by 5 to get categories from very low to very high. (44.0/5 = 8.80. We can convert this value to a round number, i. e., t 9.0

(d) Determine the number of the categories alongwith the range of each category. Add 9.0 to the lowest value of 47.0 as so on. We will finally get following categories :

47 – 56	Very low (Bihar, Jharkhand, Arunachal Pradesh, Jammu and Kashmir)
56 – 65	Low (Uttar Pradesh, Rajasthan, Andhra Pradesh, Meghalaya, Odisha, Assam, Madhya Pradesh, Chhattisgarh)

© NCERT not to be republished

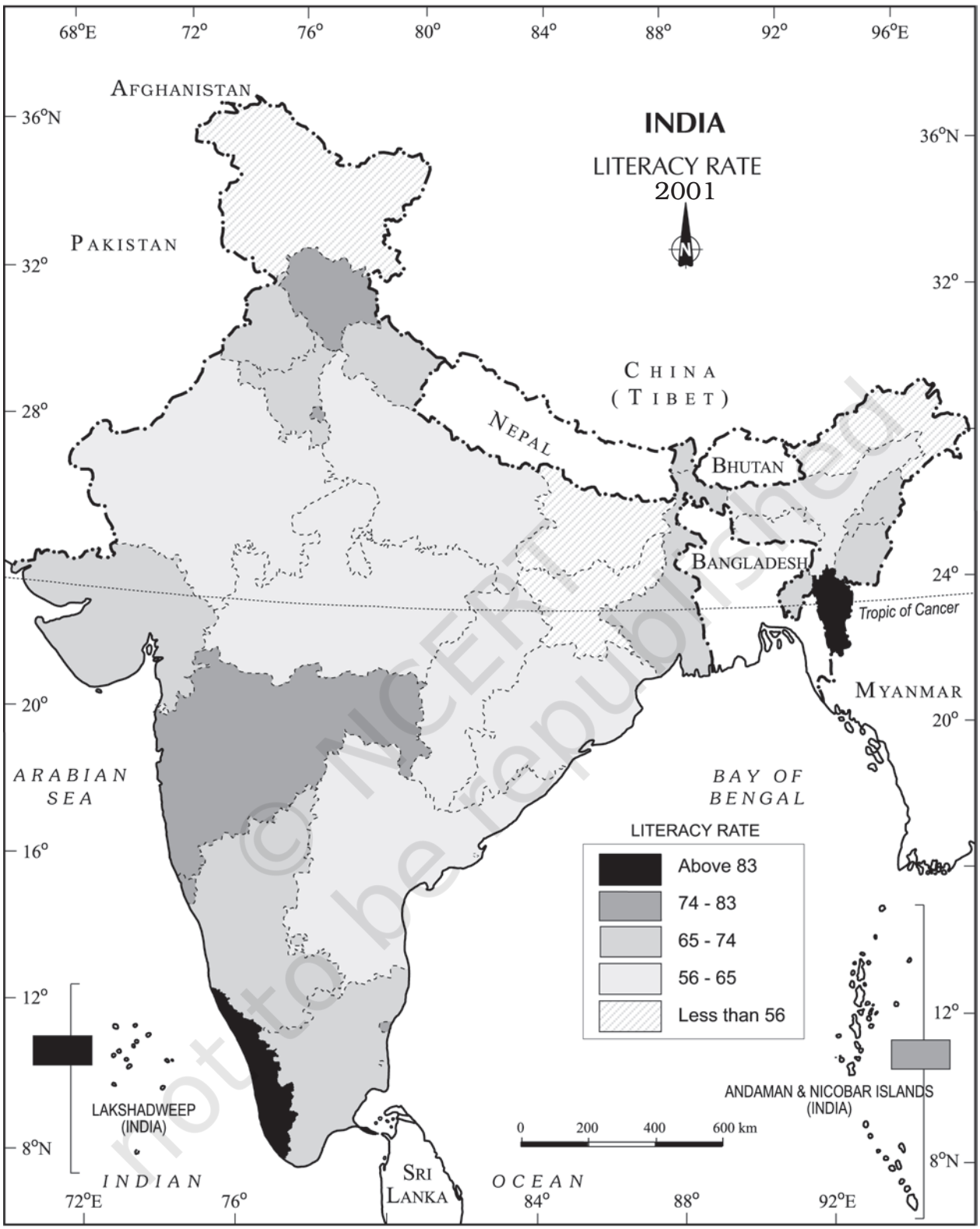


Fig. 3.14 : Literacy Rate, 2001

© NCERT not to be republished

Table 3.10 : Literacy Rate in India, 2001

Original Data on Literacy in India		
S. No.	**States / Union Territories**	**Literacy Rate**
1.	Jammu & Kashmir	55.5
2.	Himachal Pradesh	76.5
3.	Punjab	69.7
4.	Chandigarh	81.9
5.	Uttarakhand	71.6
6.	Haryana	67.9
7.	Delhi	81.7
8.	Rajasthan	60.4
9.	Uttar Pradesh	56.3
10.	Bihar	47.0
11.	Sikkim	68.8
12.	Arunachal Pradesh	54.3
13.	Nagaland	66.6
14.	Manipur	70.5
15.	Mizoram	88.8
16.	Tripura	73.2
17.	Meghalaya	62.6
18.	Assam	63.3
19.	West Bengal	68.6
20.	Jharkhand	53.6
21.	Odisha	63.1
22.	Chhattisgarh	64.7
23.	Madhya Pradesh	63.7
24.	Gujarat	69.1
25.	Daman & Diu	78.2
26.	Dadra & Nagar Haveli	57.6
27.	Maharashtra	76.9
28.	Andhra Pradesh	60.5
29.	Karnataka	66.6
30.	Goa	82.0
31.	Lakshadweep	86.7
32.	Kerala	90.9
33.	Tamil Nadu	73.5
34.	Puducherry	81.2
35.	Andaman & Nicobar Islands	81.3

Data on Literacy in India as arranged in Ascending order		
S. No.	**States / Union Territories**	**Literacy Rate**
10.	Bihar	47.0
20.	Jharkhand	53.6
12.	Arunachal Pradesh	54.3
01.	Jammu & Kashmir	55.5
9.	Uttar Pradesh	56.3
26.	Dadra & Nagar Haveli	57.6
08.	Rajasthan	60.4
28.	Andhra Pradesh	60.5
17.	Meghalaya	62.6
21.	Odisha	63.1
18.	Assam	63.3
23.	Madhya Pradesh	63.7
22.	Chhattisgarh	64.7
13.	Nagaland	66.6
29.	Karnataka	66.6
06.	Haryana	67.9
19.	West Bengal	68.6
11.	Sikkim	68.8
24.	Gujarat	69.1
03.	Punjab	69.7
14.	Manipur	70.5
05.	Uttarakhand	71.6
16.	Tripura	73.2
33.	Tamil Nadu	73.5
02.	Himachal Pradesh	76.5
27.	Maharashtra	76.9
25.	Daman & Diu	78.2
34.	Puducherry	81.2
35.	Andaman & Nicobar Islands	81.3
07.	Delhi	81.7
04.	Chandigarh	81.9
30.	Goa	82.0
31.	Lakshadweep	86.7
15.	Mizoram	88.8
32.	Kerala	90.9

65 – 74 Medium (Nagaland, Karnataka, Haryana, West Bengal, Sikkim, Gujarat, Punjab, Manipur, Uttarakhand, Tripura, Tamil Nadu)

74 – 83 High (Himachal Pradesh, Maharashtra, Delhi, Goa)

83 – 92 Very high (Mizoram, Kerala)

(e) Assign shades/pattern to each category ranging from lower to higher hues.

(f) Prepare the map as shown in *Fig. 3.14*.

(g) Complete the map with respect to the attributes of map design.

© NCERT not to be republished

Isopleth Map

We have seen that the data related to the administrative units are represented using choropleth maps. However, the variations within the data, in many cases, may also be observed on the basis of natural boundaries. For example, variations in the degrees of slope, temperature, occurrence of rainfall, etc. possess characteristics of the continuity in the data. These geographical facts may be represented by drawing the lines of equal values on a map. All such maps are termed as Isopleth Map. The word **Isopleth** is derived from **Iso** meaning equal and **pleth** means lines. Thus, an imaginary line, which joins the places of equal values, is referred as Isopleth. The more frequently drawn isopleths include Isotherm (equal temperature), Isobar (equal pressure), Isohyets (equal rainfall), Isonephs (equal cloudiness), Isohels (equal sunshine), contours (equal heights), Isobaths (equal depths), Isohaline (equal salinity), etc.

Requirement

(a) Base line map depicting point location of different places.

(b) Appropriate data of temperature, pressure, rainfall, etc. over a definite period of time.

(c) Drawing instrument specially French Curve, etc.

Rules to be observed

(a) An equal interval of values be selected.

(b) Interval of 5, 10, or 20 is supposed to be ideal.

(c) The value of Isopleth should be written along the line on either side or in the middle by breaking the line.

Interpolation

Interpolation is used to insert the intermediate values between the observed values of at two stations/locations, such as temperature recorded at Chennai and Hyderabad or the spot heights of two points. Generally, drawing of isopleths joining the places of same value is also termed as interpolation.

Method of Interpolation

For interpolation, follow the following steps:

(a) Firstly, determine the minimum and maximum values given on the map.

(b) Calculate the range of value i.e. Range = maximum value – minimum value.

(c) Based on range, determine the interval in a whole number like 5, 10, 15, etc.

The exact point of drawing an Isopleth is determined by using the following formulae.

$$\text{Point of Isopleth} = \frac{\text{Distance between two points in cm}}{\text{Difference between the two values of corresponding points}} \times \text{Interval}$$

© NCERT not to be republished

The interval is the difference between the actual value on the map and interpolated value. For example, in an Isotherm map of two places show 28 °C and 33 °C and you want to draw 30 °C isotherm, measure the distance between the two points. Suppose, the distance is 1cm or 10 mm and the difference between 28 and 33 is 5, whereas, 30 is 2 points away from 28 and 3 points behind 33, thus, exact point of 30 will be

Thus, isotherm of 30 °C will be plotted 4mm away from 28 °C or 6mm ahead of 33 °C.

(d) Draw the isopleths of minimum value first; other isopleths may be drawn accordingly.

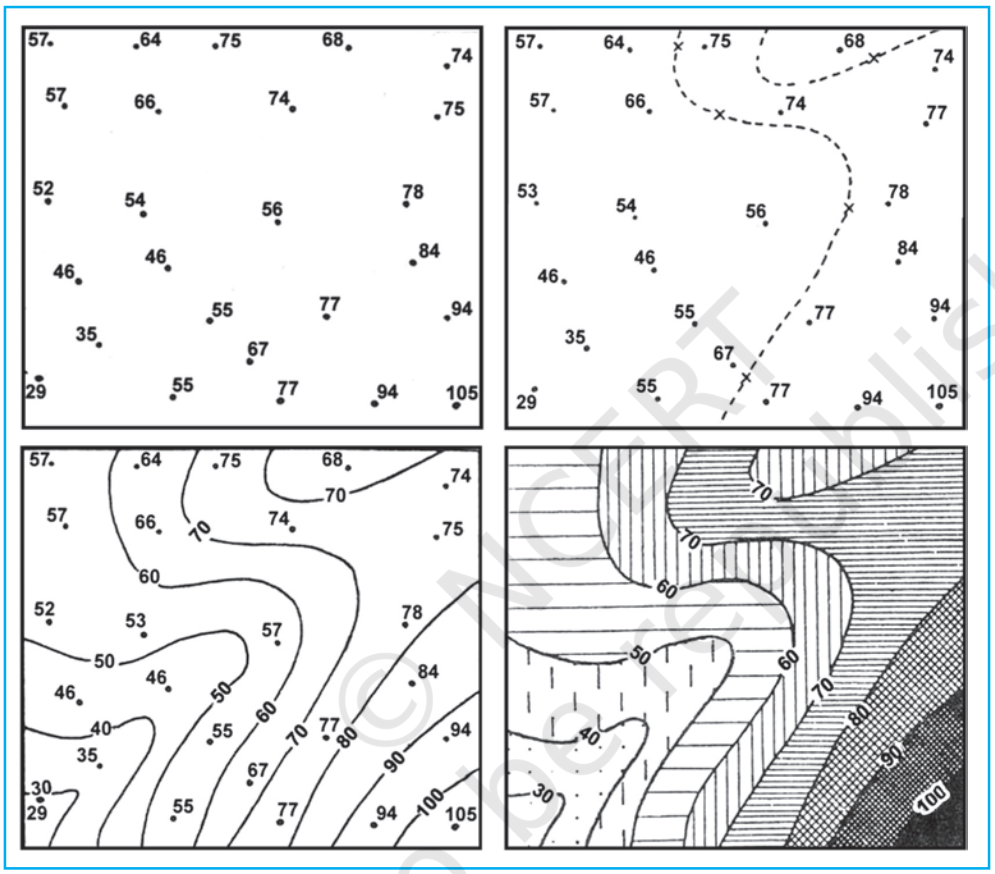


Fig. 3.15 : Drawing of Isopleths

Excercises

1. Choose the right answer from the four alternatives given below:

(i) Which one of the following map shows the population distribution:

(a) Choropleth maps (b) Isopleth maps

(c) Dot maps (d) Square root maps

(ii) Which one of the following is best suited to represent the decadal growth of population?

(a) Line graph (b) Bar diagram

(c) Circle diagram (d) Flow diagram

© NCERT not to be republished

(iii) Polygraph is constructed to represent:
(a) Only one variable (b) Two variables only
(c) More than two variables (d) None of the above

(iv) Which one of the following maps is known as "Dynamic Map"?
(a) Dot map (b) Choropleth
(c) Isopleth (d) Flow map

2. Answer the following questions in about 30 words:

(i) What is a thematic map?

(ii) Differentiate between multiple bar diagram and compound bar diagram.

(iii) What are the requirements to construct a dot map?

(iv) Describe the method of constructing a traffic flow map.

(v) What is an Isopleth map ? How an interpolation is carried out?

(vi) Describe and illustrate important steps to be followed in preparing a choropleth map.

(vii) Discuss important steps to represent data with help of a pie-diagram.

Activity

1. Represent the following data with the help of suitable diagram.

India : Trends of Urbanisation 1901-2001

Year	*Decennial growth (%)*
1911	0.35
1921	8.27
1931	19.12
1941	31.97
1951	41.42
1961	26.41
1971	38.23
1981	46.14
1991	36.47
2001	31.13

2. Represent the following data with the help of suitable diagram.

India : Literacy and Enrolment Ratio in Primary and Upper Primary Schools

Year	*Literacy Ratio*			*Enrolment Ratio Primary*			*Enrolment Ratio Upper Primary*		
	Person	Male	Female	Boys	Girls	Total	Boys	Girls	Total
1950-51	18.3	27.2	8.86	60.6	25	42.6	20.6	4.6	12.7
1999-2000	65.4	75.8	54.2	104	85	94.9	67.2	50	58.8

3. Represent the following data with help of pie-diagram.

© NCERT not to be republished

India : Land use 1951-2001

	1950-51	1998-2001
Net Sown Area	42	46
Forest	14	22
Not available for cultivation	17	14
Fallow Land	10	8
Pasture and Tree	9	5
Cultruable Waste Land	8	5

4. Study the table given below and draw the given diagrams/maps.

Area and Production of Rice in major States

States	*Area in 000 ha*	*% to total area*	*Production 000 tones*	*% to total production*
West Bengal	5,435	12.3	12,428	14.6
Uttar Pradesh	5,839	13.2	11,540	13.6
Andhra Pradesh	4,028	9.1	12,428	13.5
Punjab	2,611	5.9	9,154	10.8
Tamil Nadu	2,113	4.8	7,218	8.5
Bihar	3,671	8.3	5,417	6.4

(a) Construct a multiple bar diagram to show area under rice in each State.

(b) Construct a pie-diagram to show the percentage of area under rice in each State.

(c) Construct a dot map to show the production of rice in each State.

(d) Construct a Choropleth map to show the percentage of production of rice in States.

5. Show the following data of temperature and rainfall of Kolkata with a suitable diagram.

Months	*Temperature in º C*	*Rainfall in cm*
Jan.	19.6	1.2
Feb.	22.0	2.8
Mar.	27.1	3.4
Apr.	30.1	5.1
May	30.4	13.4
June	29.9	29.0
Jul.	28.9	33.1
Aug.	28.7	33.4
Sep.	28.9	25.3
Oct.	27.6	12.7
Nov.	23.4	2.7
Dec.	19.7	0.4

© NCERT not to be republished

12101CH06

4 Spatial Information Technology

You know that the computers enhance our capabilities in data processing and in drawing graphs, diagrams and maps. The disciplines that deals with the principles and methods of data processing and mapping using a combination of computer hardware and the application software are referred as the **Database Management System** (DBMS) and the **Computer Assisted Cartography,** respectively. However, the role of such computer applications is restricted to merely processing of the data and their graphical presentation. In other words, the data so processed or the maps and diagrams so prepared could not be used to evolve a decision support system. As a matter of fact, there are several questions that we normally encounter in our day-to-day life and look for satisfactory solutions. These questions may be: What is where ? Why is it there ? What will happen if it is shifted to a new location? Who will be benefited by such a reallocation? Who are expected to loose the benefits if reallocation takes place? In order to, understand these and many other questions, we need to capture the necessary data collected from different sources and integrate them using a computer that is supported by geo-processing tools. Herein lays the concept of a **Spatial Information System**. In the present chapter, we will discuss basic principles of the **Spatial Information Technology** and its extension to the Spatial Information System, which is more commonly known as **Geographical Information System**.

What is Spatial Information Technology?

The word **spatial** is derived from **space**. It refers to the features and the phenomena distributed over a geographically definable space, thus, having physically measurable dimensions. We know that most data that are used today have spatial components (location), such as an address of a municipal facility, or the boundaries of an agricultural holdings, etc. Hence, the Spatial Information Technology relates to the use of the technological inputs in collecting, storing, retrieving, displaying, manipulating, managing and analysing the spatial information. It is an amalgamation of Remote Sensing, GPS, GIS, Digital Cartography and Database Management Systems.

© NCERT not to be republished

What is GIS (Geographical Information System)?

The advance computing systems available since mid 1970's enable the processing of georeferenced information for the purpose of organising spatial and attribute data and their integration; locating specific information in individual files and executing the computations, performing analysis and evolving a decision support system. A system capable of all such functions is called Geographic Information System (GIS). It is defined as **A system for capturing, storing, checking, integrating, manipulating, analysing and displaying data, which are spatially referenced to the Earth. This is normally considered to involve a spatially referenced computer database and appropriate applications software.** It is an amalgamation of Computer Assisted Cartography and Database Management System and draws conceptual and methodological strength from both spatial and allied sciences such as Computer Science, Statistics, Cartography, Remote Sensing, Database Technology, Geography, Geology, Hydrology, Agriculture, Resource Management, Environmental Science, and Public Administration.

Forms of Geographical Information

Two types of the data represent the geographical information. These are spatial and non – spatial data (*Box 4.1*). The spatial data are characterised by their positional, linear and areal forms of appearances (*Fig. 4.1*).

Box 4.1 : Spatial and non-spatial data

Stock Register of a Cycle shop			Literate Population in States 1981		
Part No.	***Quantity***	***Description***	***State***	***% Male***	***% Female***
101435	54	Wheel Spoke	Kerala	75.3	65.7
108943	68	Ball Bearing	Maharashtra	58.8	34.8
105956	25	Wheel Rim	Gujarat	54.4	32.3
123545	108	Tyre	Punjab	47.2	33.7

Geographic Database : A database contains attributes and their value or class. The non-spatial data on the left display cycle parts, which can be located anywhere. The data record on the right is spatial because one of the attributes, the name of different states, which have a definite locations in a map. This data can be used in GIS.

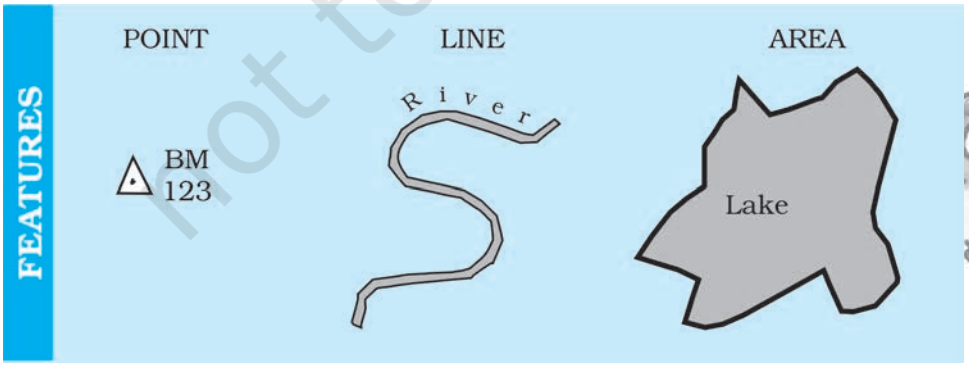


Fig. 4.1 : The Point, a Line and an Area Feature

© NCERT not to be republished

These data forms must be geometrically registered to a generally accepted and properly defined coordinate system and coded so that they can be stored in the internal database structure of GIS. On the other hand, the data those describe the spatial data are called as **Non–spatial** or attribute data. The spatial data are the most important pre-requisite in a spatial or geographical information system. In a GIS core, it could be built in several ways. These are :

- Acquire data in digital form from a data supplier
- Digitise existing analogue data
- Carry out one's own surveys of geographic entities.

The choice of a source of geographical data for a GIS application is, however, largely governed by :

- The application area in itself
- The available budget, and
- The type of data structure, i.e., vector/raster.

For many users, the most common source of spatial data is topographical or thematic maps in hard copy (paper) or soft copy form (digital). All such maps are characterised by :

- A definite scale which provides relationship between the map and the surface it represents,
- Use of symbols and colours which define attributes of entities mapped, and
- An agreed coordinate system, which defines the location of entities on the Earth's surface.

Advantages of GIS over Manual Methods

The maps, irrespective of a graphic medium of communication of geographic information and possessing geometric fidelity, are inherited with the following limitations :

(i) Map information is processed and presented in a particular way.

(ii) A map shows a single or more than one predetermined themes.

(iii) The alteration of the information depicted on the maps require a new map to be drawn.

Contrarily, a GIS possesses inherent advantages of separate data storage and presentation. It also provides options for viewing and presenting the data in several ways. The following advantages of a GIS are worth mentioning :

1. Users can interrogate displayed spatial features and retrieve associated attribute information for analysis.
2. Maps can be drawn by querying or analysing attribute data.
3. Spatial operations (polygon overlay or buffering) can be applied on integrated database to generate new sets of information.
4. Different items of attribute data can be associated with one another through shared location code.

© NCERT not to be republished

Components of GIS

The important components of a Geographical Information System include the following:

(a) Hardware (b) Software (c) Data
(d) People (e) Procedures

The different components of GIS are shown in *Fig. 4.2*.

Hardware

As discussed in Chapter 4 the GIS has three major components :

- Hardware comprising the processing, storage, display, and input and output sub-systems.
- Software modules for data entry, editing, maintenance, analysis, transformation, manipulation, data display and output.
- Database management system to take care of the data organisation.

Software

An application software with the following functional modules is important prerequisite of a GIS :

- Software related to data entry, editing and maintenance
- Software related to analysis/transformation/manipulation
- Software related to data display and output.

Data

Spatial data and related tabular data are the backbone of GIS. The existing data may be acquired from a supplier or a new data may be created/collected in-house by the user. The digital map forms the basic data input for GIS. Tabular data related to the map objects can also be attached to the digital data. A GIS will integrate spatial data with other data resources and can even use a DBMS.

People

GIS users have a wide range from hardware and software engineers to resources and environmental scientists, policy-makers, and the monitoring and implementing agencies. These cross-section of people use GIS to evolve a decision support system and solve real time problems.

Procedures

Procedures include how the data will be retrieved, input into the system, stored, managed, transformed, analysed and finally presented in a final output.

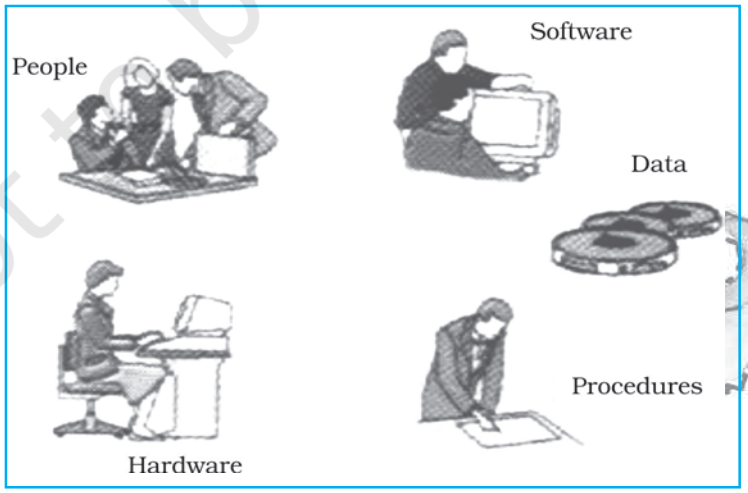


Fig. 4.2 : *Basic Components of GIS*

© NCERT not to be republished

Spatial Data Formats

The spatial data are represented in raster and vector data formats :

Raster Data Format

Raster data represent a graphic feature as a pattern of grids of squares, whereas vector data represent the object as a set of lines drawn between specific points. Consider a line drawn diagonally on a piece of paper. A raster file would represent this image by sub-dividing the paper into a matrix of small rectangles, similar to a sheet of graph paper called cells. Each cell is assigned a position in the data file and given a value based on the attribute at that position. Its row and column coordinates may identify any individual pixel (*Fig. 4.3*). This data representation allows the user to easily reconstruct or visualise the original image.

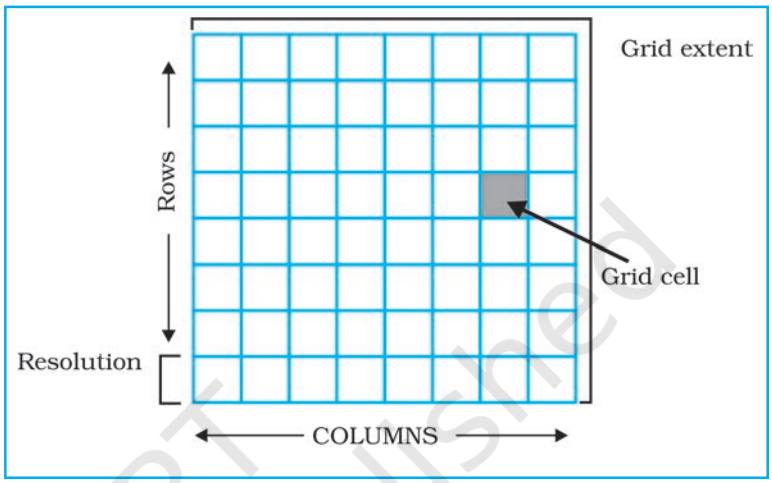


Fig. 4.3 : *Generic Structure for a Grid*

The relationship between cell size and the number of cells is expressed as the **resolution** of the raster. The effect of grid size on data in raster format is explained in *Fig. 4.4*.

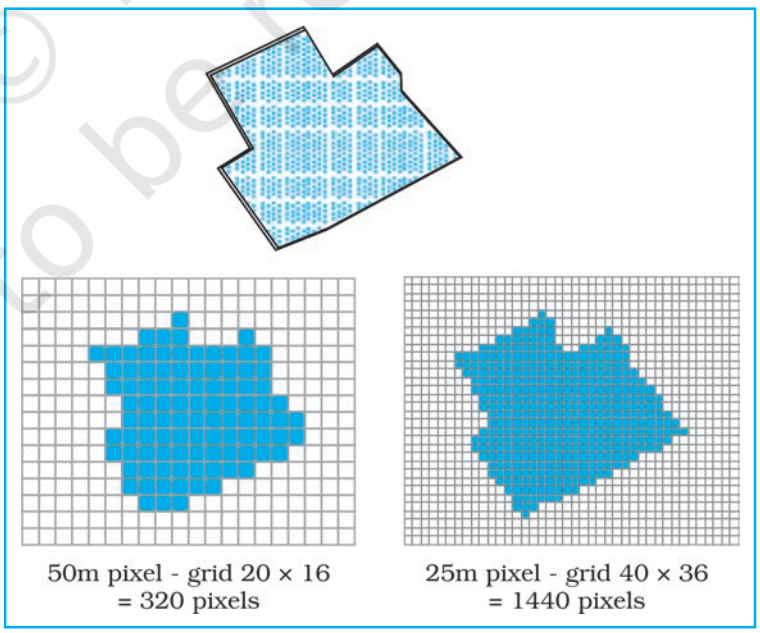


Fig. 4.4 : *Effect of Grid Size on Data in Raster Format*

© NCERT not to be republished

The Raster file formats are most often used for the following activities :

- For digital representations of aerial photographs, satellite images, scanned paper maps, etc.
- When costs need to be kept down.
- When the map does not require analysis of individual map features.
- When "backdrop" maps are required.

Vector Data Format

A vector representation of the same diagonal line would record the position of the line by simply recording the coordinates of its starting and ending points. Each point would be expressed as two or three numbers (depending on whether the representation was 2D or 3D, often referred to as X,Y or X,Y,Z coordinates) (*Fig. 4.5*). The first number, X, is the distance between the point and the left side of the paper; Y, the distance between the point and the bottom of the paper; Z, the point's elevation above or below the paper. Joining the measured points forms the vector.

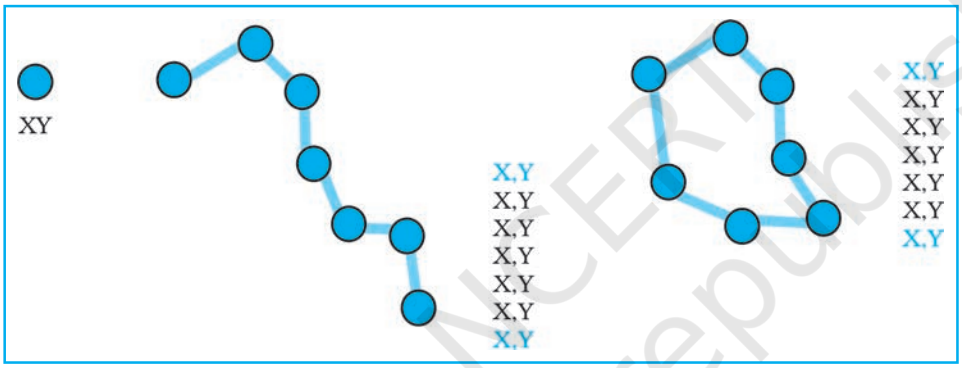


Fig. 4.5 : *The Vector Data Model is based around Coordinate Pairs*

A vector data model uses points stored by their real (earth) coordinates. Here lines and areas are built from sequences of points in order. Lines have a direction to the ordering of the points. Polygons can be built from points or lines. Vectors can store information about topology. Manual digitising is the best way of vector data input.

The Vector files are most often used for :

- Highly precise applications
- When file sizes are important
- When individual map features require analysis
- When descriptive information must be stored

The advantages and the disadvantages of the raster and vector data formats are explained in *Box 4.2*.

© NCERT not to be republished

Box 4.2 : Comparison of Raster and Vector Data Formats

Raster Model	*Vector Model*
Advantages • Simple data structure • Easy and efficient overlaying • Compatible with satellite imagery • High spatial variability is efficiently represented • Simple for own programming • Same grid cells for several attributes	**Advantages** • Compact data structure • Efficient for network analysis • Efficient projection transformation • Accurate map output
Disadvantages • Inefficient use of computer storage • Errors in perimeter and shape • Difficult network analysis • Inefficient projection transformations • Loss of information when using large cells, Less accurate (although interactive) maps	**Disadvantages** • Complex data structure • Difficult overlay operations • High spatial variability is inefficiently represented • Not compatible with satellite imagery

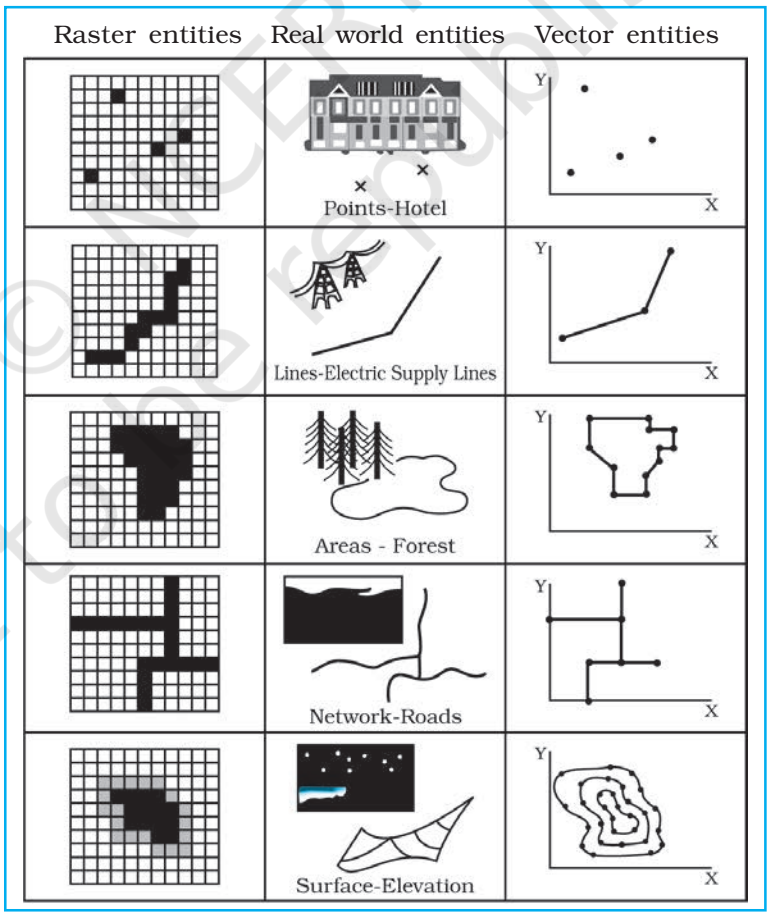


Fig. 4.6 : *Representation of Spatial Entities in Raster and Vector Data Formats*

© NCERT not to be republished

Sequence of GIS Activities

The following sequence of the activities are involved in GIS-related work :

1. Spatial data input
2. Entering of the attribute data
3. Data verification and editing
4. Spatial and attribute data linkages
5. Spatial analysis

Spatial Data Input

As already mentioned, the spatial database into a GIS can be created from a variety sources. These could be summarised into the following two categories :

(a) *Acquiring Digital Data sets from a Data Supplies*

The present day data supplies make the digital data readily available, which range from small-scale maps to the large-scale plans. For many local governments and private organisations, such data form an essential source and keep such groups of users free from overheads of digitising or collecting their own data. Although, using such existing data sets is attractive and time saving, serious attention must be paid to data compatibility when data from different sources/ supplies are combined in one project. The differences in terms of projection, scale, base level and description in attributes may cause problems.

At a practical level, users must consider the following characteristics of the data to ensure that they are compatible with the application:

- The scale of the data
- The geo-referencing system used
- The data collection techniques and sampling strategy used
- The quality of data collected
- The data classification and interpolation methods used
- The size and shape of the individual mapping units
- The length of the record.

It must also be noted that where data are used from a number of sources, and particularly where the area of study crosses administrative boundaries, the difficulties in data integration are caused by different geographical referencing systems, data classification and sampling. Hence, the user needs to be aware of these problems, which are particularly prone when compiling inter-province, and inter-district data sets. Once, the compatibility between the data acquired from different suppliers is established, the next stage involves the transfer of data from a medium of transfer to the GIS. The use of DAT tapes, CD ROMS and floppy disks is becoming increasingly common for the purpose. At this stage, the conversion from encoding and structuring system of the source to that of GIS to be used is important.

(b) *Creating digital data sets by manual input*

The manual input of data to a GIS involves four main stages :

- Entering the spatial data.
- Entering the attribute data.
- Spatial and attribute data verification and editing.
- Where necessary, linking the spatial to the attribute data.

© NCERT not to be republished

The manual data input methods depend on whether the database has a vector topology or grid cell (raster) structure. The most common ways of inputting spatial data in to a GIS are through:

- Digitisation
- Scanning

With the entity model, geographical data are in the form of points, lines and/or polygons (areas)/pixels which are defined using a series of coordinates. These are obtained by referring to the geographical referencing systems of the map or aerial photograph, or by overlaying a graticule or grid onto it. The use of digitisers and the scanners greatly reduce the time and labour involved in writing down coordinates. We shall, briefly, discuss how the spatial data are created in GIS core using a scanner.

Scanners

Scanners are the devices for converting analogue data into digital grid-based images. They are used in spatial data capture to convert a line map to high-resolution raster images which may be used directly or further processed to get vector topology. There are two basic types of scanners :

- Scanners that record data on a step-for-step basis, and
- Those that can scan whole document in one operation.

The first type of scanners incorporates a source of illumination on a movable arm (usually light emitting diodes or a stabilised fluorescent lamp) and a digital camera with high-resolution lamp. The camera is usually equipped with special sensors called Charged Coupled Devices (CCDs) arranged in an array. These are semi-conductor devices that translate the photons of light falling on their surface into counts of electrons, which are then recorded as a digital value.

The movement of either the scanner or the map builds up a digital two-dimensional image of the map. The map to be scanned can be mounted either on a flat bed, or on a rotating drum. With flatbed scanners, the light source is moved systematically up and down over the surface of the document. For large maps, scanners are used which are mounted on a stand and the illumination source and camera array are fixed in a position. The map is moved past by a feeding mechanism. Modern document scanners resemble laser printers in reverse because the scanning surface is manufactured with a given resolution of light sensitive spots that can be directly addressed by the software. There are no moving parts except a movable light source. The resolution is determined by the geometry of the sensor surface and the amount of memory rather than by a mechanical arm.

The scanned image is always far from perfect even with the best possible scanners, as it contains all the smudges and defects of the original map. The excess data, therefore, in a digital image must be removed to make it usable.

Entering the Attribute Data

Attribute data define the properties of a spatial entity that need to be handled in the GIS, but which are not spatial. For example, a road may be captured as a set of contiguous pixels or as a line entity and represented in the spatial part of the GIS by a certain colour, symbol or data location. Information describing the type of road may be included in the range of cartographic symbols. The attribute values associated with the road, such as road width, type of surface, estimated number of traffic and specific traffic regulation may also be stored separately

© NCERT not to be republished

either as spatial information in the GIS in case of relational databases, or input along with spatial description with the object-oriented data bases.

The attribute data acquired from sources like published record, official censuses, primary surveys or spread sheets can be used as input into GIS database either manually or by importing the data using a standard transfer format.

Data Verification and Editing

The spatial data captured into a GIS require verification for the error identification and corrections so as to ensure the data accuracy. The errors caused during digitisation may include data omissions, and under/over shoots. The best way to check for errors in the spatial data is to produce a computer plot or print of the data, preferably on translucent sheet, at the same scale as the original. The two maps may then be placed over each other on a light table and compared visually, working systematically from left to right and top to bottom of the map. Missing data and locational errors should be clearly marked on the printout. The errors that may arise during the capturing of spatial and attribute data may be grouped as under :

Spatial data are incomplete or double

The incompleteness in the spatial data arises through omissions in the input of points, lines, or polygons/area of manually entered data. In scanned data the omissions are usually in the form of gaps between lines where the raster vector conversion process has failed to join up all parts of a line.

Spatial data at the wrong scale

The digitising at the wrong scale produces input spatial data at a wrong scale. In scanned data, the problems usually arise during the geo-referencing process when incorrect values are used.

Spatial data are distorted

The spatial data may also be distorted if the base maps used for digitising are not scale correct. The aerial photographs, in particular, are characterised by incorrect scale because of the lens distortions, relief and till displacements. In addition, paper maps and field documents used for scanning or digitising may contain random distortions as a result of having been exposed to rain, sunshine and frequent folding. Hence, transformation from one coordinate system to another may be needed if the coordinate system of the database is different from that used in the input document or image.

These errors need corrections through various editing and updating functions as supported directly by most GIS software. The process is time-consuming and interactive that can take longer time than the data input itself. The data editing is usually undertaken by viewing the portion of map containing the errors on the computer screen and correcting them through the software using the keyboard, screen cursor controlled by a mouse or a small digitiser tablet.

Minor locational errors in a vector database may be corrected by moving the spatial entity through the screen cursor. In some GIS, computer commands may be used directly to move, rotate, erase, insert, stretch or truncate the graphical entities are required. Where excess coordinates define a line these may be removed using 'weeding' algorithms. Attribute values and spatial errors in raster data

© NCERT not to be republished

must be corrected by changing the value of the faulty cells. Once, the spatial errors have been corrected, the topology of vector line and polygon networks can be generated.

Data Conversion

While manipulating and analysing data, the same format should be used for all data. When different layers are to be used simultaneously, they should all be in vector or all in raster format. Usually, the conversion is from vector to raster, because the biggest part of the analysis is done in the raster domain. Vector data are transformed to raster data by overlaying a grid with a user-defined cell size.

Sometimes, the data in the raster format are converted into vector format. This is the case especially if one wants to achieve data reduction because the data storage needed for raster data are much larger than for vector data.

Geographic Data : Linkages and Matching

The linkages of spatial and the attribute data are important in GIS. It must, therefore, carefully be undertaken. Linking of attribute data with a non-related spatial data shall lead to chaos in ultimate data analysis. Similarly, matching of one data layer with another is also significant.

Linkages

A GIS typically links different data sets. Suppose, we want to know the mortality rate due to malnutrition among children under 10 years of age in any state. If we have one file that contains the number of children in this age group, and another that contains the mortality rate from malnutrition, we must first combine or link the two data files. Once this is done, we can divide one figure by the other to obtain the desired answer.

Exact Matching

Exact matching means when we have information in one computer file about many geographic features (e.g., towns) and additional information in another file about the same set of features. The operation to bring them together may easily be achieved using a key common to both files, i. e. name of the towns. Thus, the record in each file with the same town name is extracted, and the two are joined and stored in another file.

Hierarchical Matching

Some types of information, however, are collected in more detail and less frequently than other types of information. For example, land use data covering a large area are collected quite frequently. On the other hand, land transformation data are collected in small areas but at less frequent intervals. If the smaller areas adjust within the larger ones, then the way to make the data match of the same area is to use hierarchical matching — add the data for the small areas together until the grouped areas match the bigger ones and then match them exactly.

Fuzzy Matching

On many occasions, the boundaries of the smaller areas do not match with those of the larger ones. The problem occurs more often when the environmental data are involved. For example, crop boundaries that are usually defined by field edges/boundaries rarely match with the boundaries of the soil types. If we want

© NCERT not to be republished

to determine the most productive soil for a particular crop, we need to overlay the two sets and compute crop productivity for each soil type. This is like laying one map over another and noting the combinations of soil and productivity.

A GIS can carry out all these operations. However, the sets of spatial information are linked only when they relate to the same geographical area.

Spatial Analysis

The strength of the GIS lies in its analytical capabilities. What distinguish the GIS from other information systems are its spatial analysis functions. The analysis functions use the spatial and non-spatial attributes in the database to answer questions about the real world. Geographic analysis facilitates the study of real-world processes by developing and applying models. Such models provide the underlying trends in geographic data and thus, make new possibilities available. The objective of geographic analysis is to transform data into useful information to satisfy the requirements of the decision-makers. For example, GIS may effectively be used to predict future trends over space and time related to variety of phenomena. However, before undertaking any GIS based analysis, one needs to identify the problem and define purpose of the analysis. It requires step – by – step procedures to arrive at the conclusions. The following spatial analysis operation may be undertaken using GIS :

(i) Overlay analysis (ii) Buffer analysis

(iii) Network analysis (iv) Digital Terrain Model

However, under the constraints of time and space only the overlay and buffer analysis operations will be dealt herewith.

Overlay Analysis Operations

The hallmark of GIS is overlay operations. An integration of multiple layers of maps using overlay operations is an important analysis function. In other words, GIS makes it possible to overlay two or more thematic layers of maps of the same area to obtain a new map layer (*Fig. 4.7*). The overlay operations of a GIS are

Overlay Operation : x + y = z

where

x = Road map
y = Rail map
z = Transportation Map
+ = The spatial overlay operation 'union'

Graphical representation

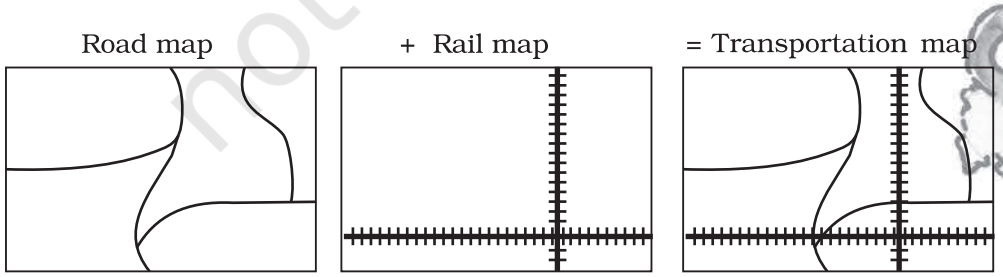


Fig. 4.7 : *Simple Overlay Operation*

© NCERT not to be republished

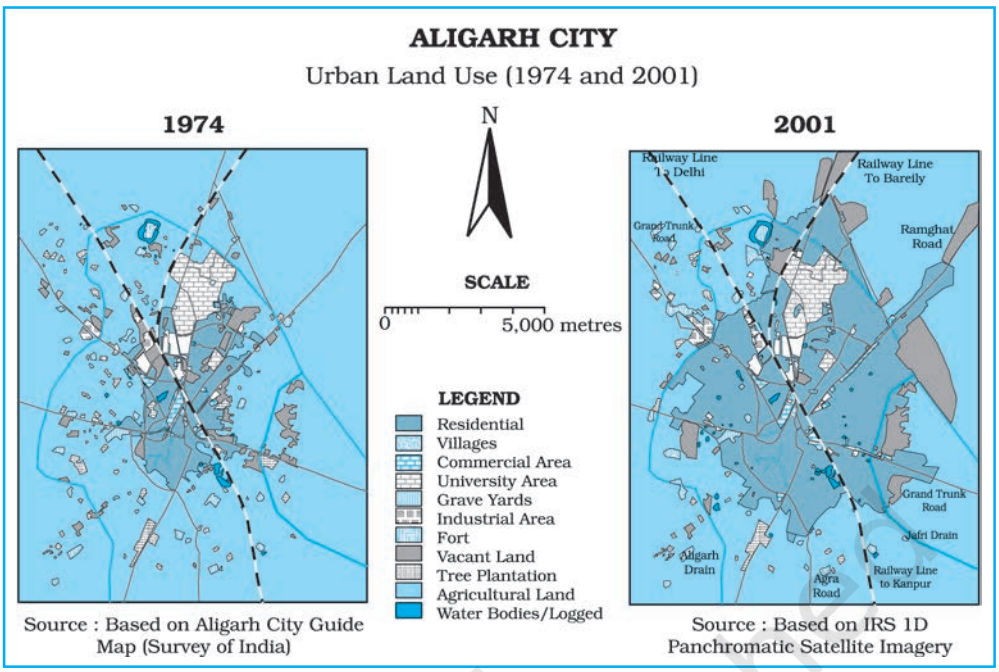


Fig. 4.8 : *Urban Land Use in Aligarh City, Uttar Pradesh during 1974 and 2001*

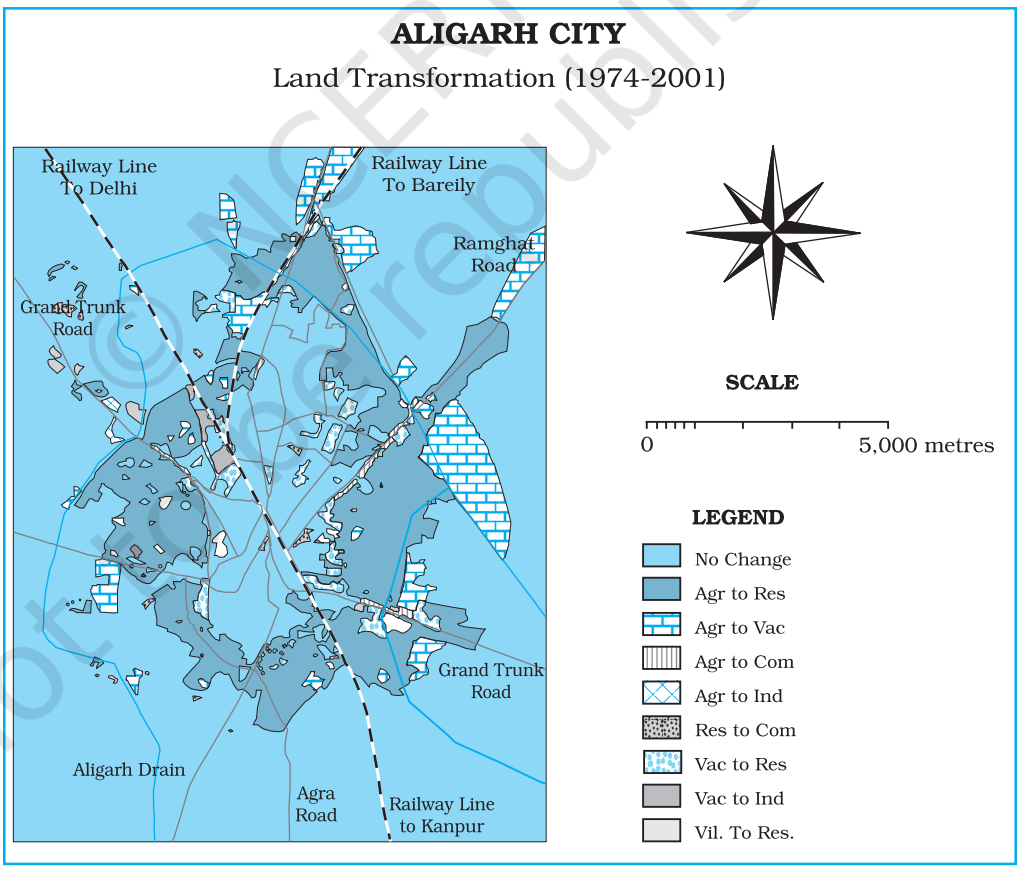


Fig. 4.9 : *Urban Land Transformations in Aligarh City during 1974-2001*

© NCERT not to be republished

similar to the sieve mapping, i.e., the overlaying of tracing of maps on a light table to make comparisons and obtain an output map.

Map overlay has many applications. It can be used to study the changes in land use/land cover over two different periods in time and analyse the land transformations. For example, *Fig. 4.8* depicts urban land use during 1974 and 2001. When the two maps overlaid, the changes in urban land use have been obtained (*Fig. 4.9*) and the urban sprawl is mapped during the given time period (*Fig. 4.10*). Similarly, overlay analysis is also useful in suitability analysis of the given land use for proposed land uses.

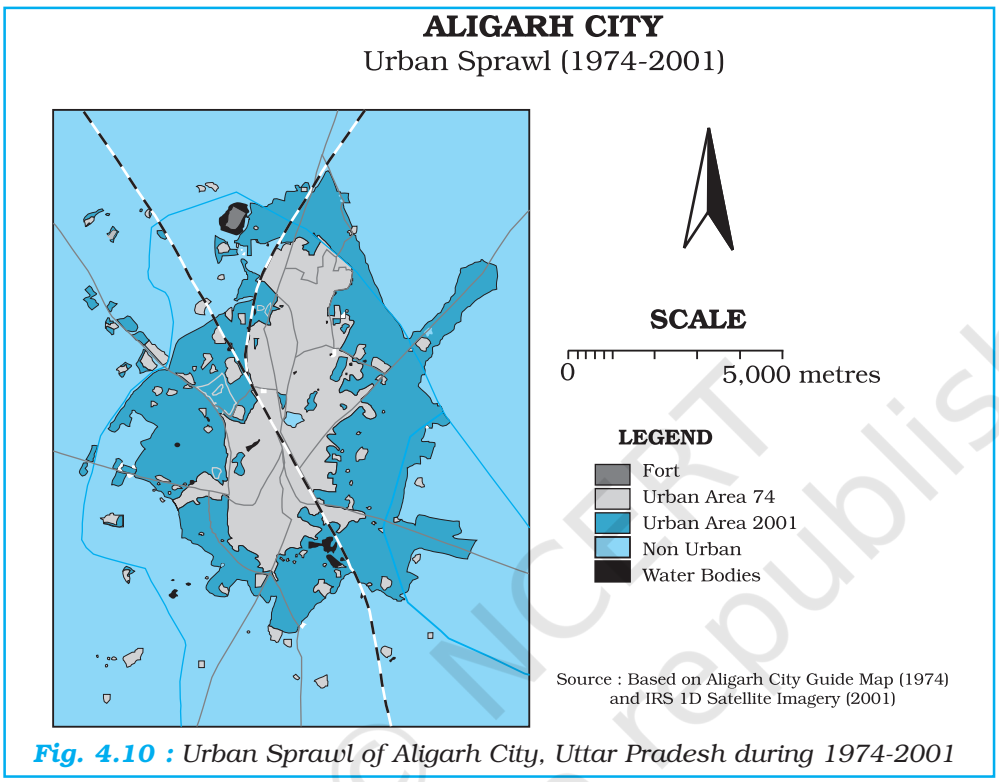


Fig. 4.10 : *Urban Sprawl of Aligarh City, Uttar Pradesh during 1974-2001*

Buffer Operation

Buffer operation is another important spatial analysis function in GIS. A buffer of a certain specified distance can be created along any point, line or area feature (*Fig. 4.11*). It is useful in locating the areas/population benefitted or denied of the facilities and services, such as hospitals, medical stores, post office, asphalt roads, regional parks, etc. Similarly, it can also be used to study the impact of point sources of air, noise or water pollution on human health and the size of the population so affected. This kind of analysis is called proximity analysis. The buffer operation will generate polygon feature types irrespective of geographic features and delineates spatial proximity. For example, numbers of household living within one-kilometre buffer from a chemical industrial unit are affected by industrial waste discharged from the unit.

Arc View/ArcGIS, Geomedia Quantum GIS free opensoftware and all other GIS softwares provide modules for buffer analysis along point, line and area features. For example, by using appropriate commands of either of the available software, one can create buffers of 2, 4, 6, 8 and 10 kilometres around the cities having a major hospital located. As a case study, point location of Saharanpur,

© NCERT not to be republished

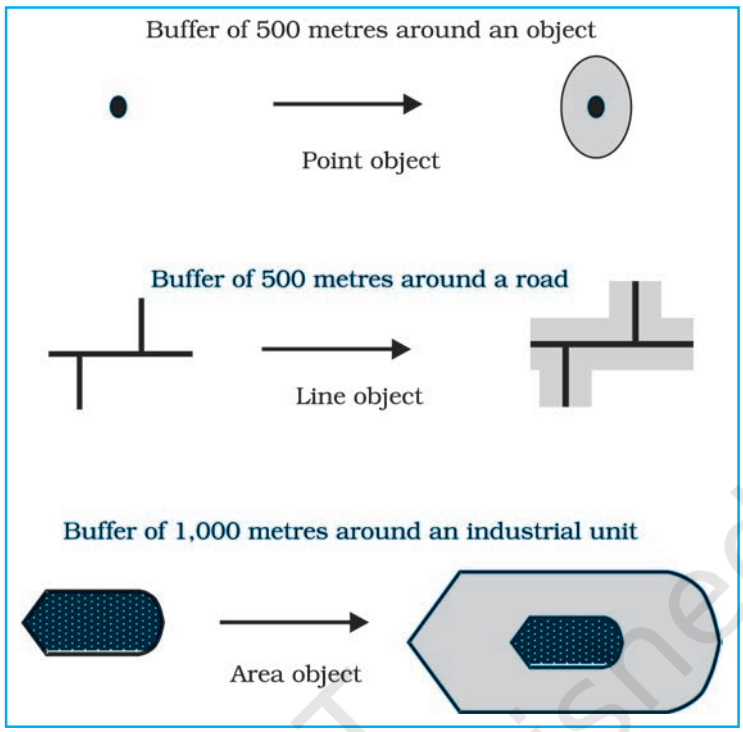


Fig. 4.11 : *Buffers of Constant Width Drawn around a Point, Line and a Polygon*

Muzaffarnagar, Meerut, Ghaziabad, Gautam Budh Nagar and Aligarh has been mapped (*Fig. 4.12*) and the buffer have been created from the cities where major hospitals are found. One can observe that the areas closer to the cities are better served, people living away from the cities have to travel long distances to utilise the medical services and their areas that are least benefitted (*Fig. 4.13*).

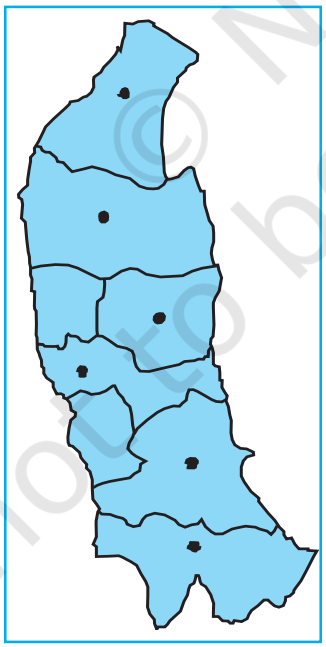


Fig. 4.12 : *Location Map of the Cities of Western Uttar Pradesh*

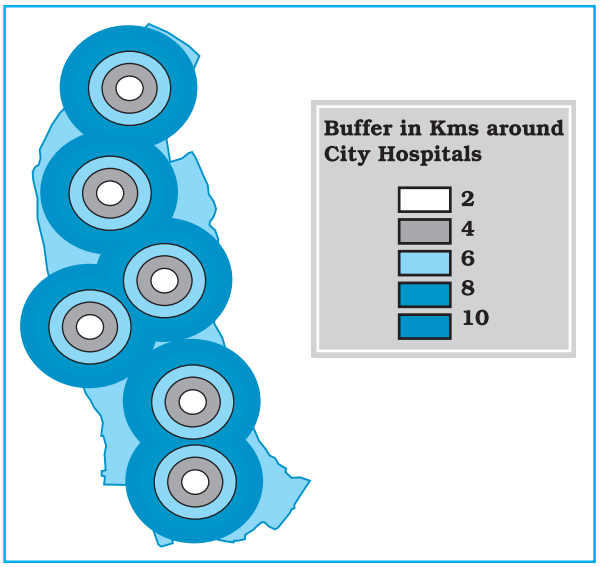


Fig. 4.13 : *Buffers of Specified Distances around Hospitals*

Internet sources to learn more:

- schoolgis.nic.in
- bhuvan.nrsc.gov.in
- www.iirs.gov.in

© NCERT not to be republished

Excercises

1. Choose the right answer from the four alternatives given below :

(i) The spatial data are characterised by the following forms of appearance :
(a) Positional (b) Linear
(c) Areal (d) All the above forms

(ii) Which one of the following operations requires analysis module software?
(a) Data storage (b) Data display
(c) Data output (d) Buffering

(iii) Which one of the following is disadvantage of Raster data format ?
(a) Simple data structure
(b) Easy and efficient overlaying
(c) Compatible with remote sensing imagery
(d) Difficult network analysis

(iv) Which one of the following is an advantage of Vector data format ?
(a) Complex data structure
(b) Difficult overlay operations
(c) Lack of compatibility with remote sensing data
(d) Compact data structure

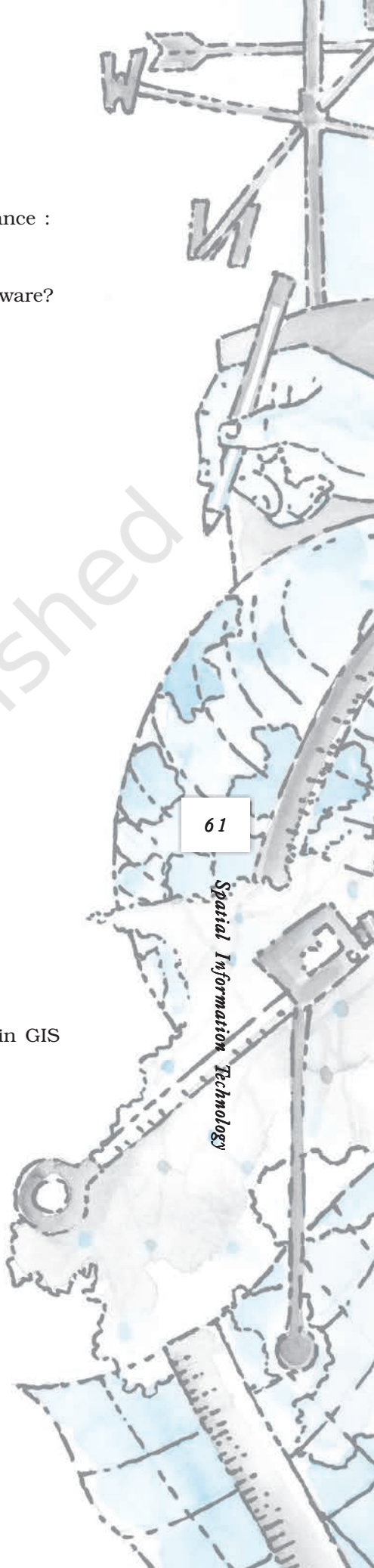
(v) Urban change detection is effectively undertaken in GIS core using:
(a) Overlay operations
(b) Proximity analysis
(c) Network analysis
(d) Buffering

2. Answer the following questions in about 30 words :

(i) Differentiate between raster and vector data models.
(ii) What is an overlay analysis?
(iii) What are the advantages of GIS over manual methods?
(iv) What are important components of GIS?
(v) What are different ways in which spatial data is built in GIS core?
(vi) What is Spatial Information Technology?

3. Answer the following questions in about 125 words :

(i) Discuss raster and vector data formats. Give example.
(ii) Write an explanatory account of the sequence of activities involved in GIS related work.



© NCERT not to be republished

Glossary

Bar Graph : A series of columns or bars drawn proportional in length to the quantities they represent. They are drawn on a selected scale. They may be drawn either horizontally or vertically.

Central Tendency : The tendency of quantitative data to cluster around some value.

Choropleth Maps : Maps drawn on quantitative areal basis, calculated as average values per unit of area within specific administrative units, e.g. density of population and percentage of urban to total population. Distribution of a given phenomenon is shown by various shades of a colour or intensity.

Class Intervals : The difference between the lower and upper limits of any class of a frequency distribution is known as its class interval.

Correlation Co-efficient : A measure of the degree and direction of relationship between two variables.

Cumulative Frequency : The measurement of distribution of values in the different class intervals expressed as a percentage of the total frequencies either above or below specified value.

Dispersion : The degree of internal variations in the different values of a variable.

Flow Maps : Maps in which the "flow" or movement of people or commodities is represented by riband whose thickness is proportional to the quantity of goods or the number of people moving along different routes.

Histogram : A graphical representation of a frequency distribution, such as seasonal frequencies of rainfall.

Mean Deviation : A measure of dispersion derives from the average of deviations from some central value. Such deviations are taken absolutely, i.e., their signs are ignored. The central value is generally mean or median.

Median : It is the value which divides the number of observations in such a way that half the value are less than this value and half of them are more. If the values of a variable are arranged in either ascending of descending order, the median is the middle value.

Mode : The mode is that value of a variable which occurs maximum number of times.

Pie Diagram : A circular diagram in which a circle is divided into sectors for presenting data in percentage.

Standard Deviation : The most commonly used measure of dispersion. The standard deviation is the positive square root of the mean of the squares of deviations from the mean.

Tabulation : The process of putting raw data into a systematically arranged tabular form.

Variable : Any characteristic which varies. A quantitative variable is a characteristic which has different values; the differences of which are quantitatively measurable. Rainfall, for example, is a quantitative variables, because the differences in its different values at different places or at different times are quantitatively measurable. A qualitative variable on the other hand, is the characteristic; the different values of which cannot be measured quantitatively. Sex, for example, is a qualitative variable, it can be either male or female. A qualitative variable is also known as an attribute.

© NCERT not to be republished

© NCERT not to be republished

Notes

© NCERT not to be republished